Stress-based Finite Element Methods for Variational Inequalities in Contact Mechanics

Von der Fakultät für Mathematik der Universität Duisburg-Essen
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation von

Bernhard Kober, M.Sc.

Erstgutachter: Prof. Dr. Gerhard Starke, Universität Duisburg-Essen
Zweitgutachter: Prof. Dr. Rolf Krause, Università della Svizzera italiana

Ort und Datum der Einreichung: Essen, am 05.11.19
Ort und Datum der Disputation: Essen, am 23.01.20

DOI: 10.17185/duepublico/71265
URN: urn:nbn:de:hbz:464-20200131-102447-0

Für Maili

Abstract

A stress-based mixed finite element method with reduced symmetry for Linear Elasticity featuring Raviart-Thomas elements is extended first to the frictionless Signorini problem and then to contact problems obeying Coulomb's friction law. Different possible discretizations of the contact constraints are examined. The resulting (quasi-)variational inequalities with nonsmooth constraints are solved using Lagrange multipliers, a semismooth Newton method and, in case of Coulomb friction, a fixed point iteration. A reconstruction-based a-posteriori error estimator is derived. Its reliabilty is shown under certain regularity assumptions on the solution that correspond to a uniqueness criterion for the solution of Coulomb's problem. The efficiency of the resulting adaptive refinement strategy is tested by a number of numerical experiments in two and three dimensions.

Keywords: Contact, Coulomb Friction, Linear Elasticity, Weak Symmetry, Variational Inequalities, A-posteriori Error Estimation, Mixed Finite Element Method, Raviart-Thomas-elements

Zusammenfassung

Eine spannungsbasierte Finite Elemente Methode mit schwacher Symmetrie für Lineare Elastizität wird zuerst auf das reibungsfreie Signorini Problem und in weiterer Folge auf Kontaktprobleme mit Coulomb Reibung erweitert. Verschiedene mögliche Diskretisierungen der Kontaktbedingungen werden untersucht. Die resultierenden (Quasi-)Variationsungleichungen werden unter Zuhilfenahme von Lagrange-Multiplikatoren mit einem halbglatten Newtonverfahren und, im Fall von Coulomb Reibung, mit einer Fixpunktiteration gelöst. Ein rekonstruktionsbasierter a-posteriori Fehlerschätzer wird hergeleitet. Seine Zuverlässigkeit wird unter bestimmten Regularitätsvorraussetzungen an die Lösung gezeigt, welche in Zusammenhang mit einem Eindeutigkeitskriterium für die Lösung des Coulomb Problems stehen. Die Effizienz der resultierenden adaptiven Verfeinerungsstrategie wird an einigen numerischen Experimenten in zwei und drei Raumdimensionen getestet.

Schlagwörter: Kontakt, Reibung, Lineare Elastizität, Schwache Symmetrie, Variationsungleichungen, A-posteriori Fehlerschätzung, Gemischte Finite Elemente Methode, Raviart-Thomas-Elemente

Contents

Introduction

This work is about the numerical analysis of stress-based adaptive finite element methods for the solution of variational inequalities arising from contact problems in solid mechanics.

Contact problems in solid mechanics deal with the deformation and resultant internal stresses of solids that are subject both to known loads and a-priori unknown forces arising from the contact with other solids. They arise in many applications, not only in engineering but also in medicine (cf. [KKS$^+$08]). While a realistic model of contact includes frictional effects, the situation is sometimes idealized by neglecting friction in order to make the problem more manageable. Such an idealization leads to the well-known problem of Signorini ([Sig33]). In this work we will consider the interaction of only one elastic solid with a rigid obstacle. Both the frictionless and frictional case will be examined.

The mathematical formulation of such contact models typically results in optimization problems with convex admissible sets, yielding variational inequalities as optimality conditions which can be solved using finite element methods. The theory of variational inequalites goes back to [LS67] and has been applied to fluid and solid mechanics in [KS80] and to contact problems in [KO88] and [HHNL88]. Most of the results in these fundamental works (and consequently in the literature in general) focus on the formulation of the problem in displacements, which is sometimes called the primal approach. However, the stress distribution is often what is of higher interest in applications (cf. [Bra13, HHN96]). Furthermore many phenomena such as friction or plasticity are modeled mainly in terms of stress and conservation of momentum is typically improved by stress-based approaches. The powerful concept of duality which can be employed for the derivation of stress-based variational formulations is already noted in some of the classical works mentioned above. It has also been studied in depth more recently in [Cap14] in combination with a regularized nonlocal friction law.

A major challenge in the design of finite element methods for stress-based approaches lies in the conservation of angular momentum expressed by the symmetry of the stress-tensor. Since methods that impose strong symmetry are complicated and computationally expensive (cf. [AW02]), there has been a growing interest in relaxing the symmetry constraint (e.g. [ABD84, BBF09]). In this work we want to apply the ideas of finite element discretization featuring weak symmetry to frictional contact problems.

Even with relaxed symmetry, discretizing the stress field leads to a large number of

degrees of freedom. This challenge together with the facts that contact is typically a local phenomenon and elasticity problems often have singular solutions call for the use of adaptive mesh refinement strategies. For this purpose reliable (and if possible efficient) a-posteriori error estimation is needed. Residual-based a-posteriori error estimators for the primal problem have been proposed for the frictionless Signorini problem in [KVW15] and for the problem with Coulomb friction in [HL09]. Furthermore a least squares formulation, which typically includes an a-posteriori error estimator, of the frictionless Signorini problem is analyzed in [ACS09] and [KMS17]. According to our knowledge the topic of a-posteriori error estimation for the dual formulation of frictional contact has not yet been treated in the literature. Therefore a main goal of this thesis is the derivation of a reliable reconstruction-based a-posteriori error estimator for frictional contact problems.

In the first chapter of this work we summarize a number of fundamental definitions and results needed in the subsequent chapters. We start with a short introduction of Sobolev spaces and the definition of the subspaces and subsets used for the treatment of contact problems. Particular focus will be put on the somewhat delicate definition of trace spaces. Next we review the needed finite element spaces and their approximation properties and conclude with a collection of useful inequalities for elasticity problems.

The main part of this work consists of the chapters 2-4. We start with the problem of linear elasticity and gradually extend it, first to frictionless contact and then to contact with Coulomb friction. In chapter 2 we review the stress-based mixed formulation known as the generalized Hellinger-Reissner principle. The relationship of this formulation to the standard formulation is analysed in detail, since it is later used in the derivation of a reliable a-posteriori error estimate. The finite element discretization that will be used throughout this work as well as a displacement reconstruction procedure are introduced. Finally we present a short numerical example to illustrate the behaviour of the adaptive refinement strategy resulting from the use of the proposed error estimator.

In chapter 3 we consider the Signorini problem and extend the variational formulation to respect the contact conditions and preserve strong duality. Different discretization approaches for the sign condition are discussed and the corresponding a-priori and a-posteriori error analysis is performed. The semi-smooth Newton method is introduced and the performance of the extended adaptive finite element method is tested on a number of numerical experiments.

Chapter 4 deals with the challenges arising from incorporating the Coulomb friction law in the contact model. An auxiliary problem with given friction is introduced and used for a sound formulation of the Coulomb problem. The discretization of the friction condition and the application of the semi-smooth Newton method to the frictional problem are discussed. The extension of the reliability result of the proposed a-posteriori error estimate is obtained under certain regularity assumptions on the solution. The chapter concludes with the presentation of several numerical examples in two and three dimensions.

We conclude this thesis with a summary of the main results and an outlook considering some open questions that arose in the course of this work.

Chapter 1

Preliminaries

1.1 Spaces and Traces

In this section we will define the Hilbert spaces that are used in this work and recapitulate a number of fundamental results related to those spaces. Finally we will specify a number of subspaces that will be useful in the formulation of contact problems in solid mechanics.

1.1.1 Sobolev Spaces

Sobolev spaces are fundamental to the theory of weak solutions of partial differential equations. While an elaborate examination of the rich theory of Sobolev spaces can be found in books such as [AF03, LM72, Gri85] we will only present a minimal excerpt.

For this purpose let $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ be a bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$. Let furthermore $\Gamma_* \subset \Gamma$ be an open subset of the boundary. The sobolev spaces $H^m(\Omega)$ are usually defined as spaces of weakly differentiable functions in the following way for non-negative integers m:

$$H^m(\Omega) := \{v \in L^2(\Omega) : D^\alpha v \in L^2(\Omega) \,\forall\, |\alpha| \leq m\} \tag{1.1}$$

It can be shown that this definition is equivalent to taking the closure of $H^m(\Omega) \cap C^\infty(\Omega)$ with respect to the norm induced by the following scalar product:

$$(u, v)_{H^m(\Omega)} := \sum_{|\alpha| \leq m} (D^\alpha u, D^\alpha v)_{L^2(\Omega)} \tag{1.2}$$

It is evident from this equivalence that $H^m(\Omega)$ is a Hilbert space and we will subsequently abbreviate the corresponding scalar products and norms with

$$\begin{aligned} (u, v)_{m,\Omega} &:= (u, v)_{H^m(\Omega)}\,, \\ \|v\|_{m,\Omega} &:= \sqrt{(v, v)_{H^m(\Omega)}}\,, \end{aligned}$$

omitting the subscript Ω whenever it will cause no confusion and dropping the subscript altogether whenever it is clear that we are operating in $H^0(\Omega) = L^2(\Omega)$.

This concept is generalized to Sobolev spaces with real exponents $H^{m+s}(\Omega)$ with $s \in (0,1)$, either by taking the closure with respect to "intermediate" norms

$$\|v\|^2_{m+s,\Omega} = \|v\|^2_{m,\Omega} + \sum_{|\alpha|=m} \iint_{\Omega\,\Omega} \frac{|D^\alpha v(x) - D^\alpha v(y)|^2}{|x-y|^{d+2s}}\, dx\, dy \tag{1.3}$$

or by interpolation between the spaces $H^m(\Omega)$ and $H^{m+1}(\Omega)$. The resulting spaces and norms are equivalent and we have the inclusion $H^s(\Omega) \subset H^t(\Omega)$ for $s \geq t$ (see [BS08, Chapter 14]).

By $H^s_0(\Omega)$ we denote the closure of the space of smooth functions with compact support $C^\infty_c(\Omega)$ with respect to $\|\cdot\|_s$. Furthermore $H^1_{\Gamma_*}(\Omega)$ is defined as the closure of the space of functions $v \in C^\infty(\overline{\Omega}) \cap H^1(\Omega)$ that vanish on Γ_*. Then we have

$$H^1_0(\Omega) \subset H^1_{\Gamma_*}(\Omega) \subset H^1(\Omega)\,. \tag{1.4}$$

We point out that for domains with smooth boundary we have

$$H^s_0(\Omega) = H^s(\Omega) \tag{1.5}$$

if and only if $s \leq \frac{1}{2}$ (see [LM72, Theorem 11.1]). The case of $s = m + \frac{1}{2}$ is particularly delicate. While in all other cases a function in $H^s_0(\Omega)$ does allow an extension by zero outside of Ω, this is not the case for $H^{m+1/2}_0(\Omega)$ (see [LM72, Theorem 11.4]). However, it is possible to define subspaces, whose elements allow such an extension. For our purposes we will need the space $H^{1/2}_{00}(\Omega)$, which can be characterized by

$$H^{1/2}_{00}(\Omega) := \{v \in H^{1/2}(\Omega) : v \text{ allows an extention by zero outside of } \Omega\}$$

and whose rather technical definition can be found in [LM72, Theorem 11.7] or [KO88, Chapter 5].

Finally we introduce the Hilbert space $H(\mathrm{div}, \Omega)$ which is especially important for mixed problems and which is defined as

$$H(\mathrm{div}, \Omega) := \{\mathbf{v} \in L^2(\Omega)^d : \mathrm{div}\, \mathbf{v} \in L^2(\Omega)\} \tag{1.6}$$

with corresponding scalar product and norm

$$\begin{aligned} (\mathbf{u}, \mathbf{v})_{\mathrm{div},\Omega} &:= (\mathbf{u}, \mathbf{v})_0 + (\mathrm{div}\, \mathbf{u}, \mathrm{div}\, \mathbf{v})_0\,, \\ \|\mathbf{v}\|^2_{\mathrm{div},\Omega} &:= \|\mathbf{v}\|^2_0 + \|\mathrm{div}\, \mathbf{v}\|^2_0\,. \end{aligned}$$

Here we also introduced the notation of bold symbols indicating d-dimensional vectors. We will use this notation also for the corresponding function spaces e.g. $\mathbf{H}^1(\Omega) = H^1(\Omega)^d$. Elements of $\mathbf{H}(\mathrm{div}, \Omega) = H(\mathrm{div}, \Omega)^d$ which are $d \times d$-dimensional tensors will also be written in bold.

1.1.2 Traces of functions in Sobolev Spaces

Since the phenomenon of contact first and foremost affects the surface of a solid, much of this work will be concerned with activities on the boundary Γ. Thus, we need to introduce some related notation and review some properties of the boundary values of the functions in the Sobolev spaces which we introduced in the previous section. We start with the well known trace theorem for $H^1(\Omega)$.

Theorem 1.1. *The trace operator $\gamma(v) := v|_\Gamma$ defined on $C^0(\overline{\Omega})$ has an continuous extention to $H^1(\Omega)$ that satisfies*

$$\gamma(H^1(\Omega)) = H^{1/2}(\Gamma)\,, \tag{1.7}$$

$$\|\gamma(v)\|_{1/2,\Gamma} \lesssim \|v\|_{1,\Omega}\,. \tag{1.8}$$

Proof. See [EG04, Theorem B.52] and [Neč12, Chapter 2, Theorem 5.5]. □

(Here and throughout this work the symbol $\lesssim$ is used to denote that a generic constant is part of the inequality.) Furthermore for any given $v \in H^{1/2}(\Gamma)$ there exists a uniquely determined $\bar{v} \in H^1(\Omega)$ with

$$\|\gamma(\bar{v})\|_{1/2,\Gamma} = \|\bar{v}\|_{1,\Omega} \tag{1.9}$$

and $\gamma(\bar{v}) = v$ (see [BBF13, Section 2.1.1]).

Next, we are interested in the dual space of $H^{1/2}(\Gamma)$ which is denoted by $H^{-1/2}(\Gamma)$ and equipped with the dual norm

$$\|v\|_{-1/2,\Gamma} = \sup_{u\in H^{1/2}(\Gamma)} \frac{\langle v, u\rangle_\Gamma}{\|u\|_{1/2,\Gamma}} \tag{1.10}$$

where $\langle\cdot,\cdot\rangle_\Gamma$ denotes the duality pairing on $H^{-1/2}(\Gamma) \times H^{1/2}(\Gamma)$ for which

$$\langle v, u\rangle_\Gamma = (v, u)_{0,\Gamma} \tag{1.11}$$

holds, whenever $v \in L^2(\Gamma)$. Green's formula

$$\int_\Gamma u\mathbf{v}\cdot\mathbf{n}\,dx = \int_\Omega \mathbf{v}\nabla u\,dx + \int_\Omega u\operatorname{div}\mathbf{v}\,dx\,, \tag{1.12}$$

with $\mathbf{n}$ denoting the outer unit normal, suggests that for $\mathbf{v} \in H(\operatorname{div},\Omega)$ we can define a normal trace in $H^{-1/2}(\Gamma)$. Indeed we have the following trace theorem for $H(\operatorname{div},\Omega)$:

Theorem 1.2. *There exists a uniquely determined linear continuous mapping (which we call the normal trace operator) $\pi : H(\operatorname{div},\Omega) \to H^{-1/2}(\Gamma)$ that satisfies*

$$\pi(H(\operatorname{div},\Omega)) = H^{-1/2}(\Gamma)\,, \tag{1.13}$$

$$\|\pi(\mathbf{v})\|_{-1/2,\Gamma} \lesssim \|\mathbf{v}\|_{\operatorname{div},\Omega}\,, \tag{1.14}$$

as well as $\pi(\mathbf{v}) = (\mathbf{v}\cdot\mathbf{n})|_\Gamma$ whenever $\mathbf{v}$ is smooth enough for the right hand side to be defined.

Proof. See [GR86, Chapter 2, Theorem 2.5] and [BBF13, Section 2.1.1]. □

Again there exists a $\bar{\mathbf{v}} \in H(\mathrm{div}, \Omega)$ for any given $v \in H^{-1/2}(\Gamma)$ with

$$\|\pi(\bar{\mathbf{v}})\|_{-1/2,\Gamma} = \|\bar{\mathbf{v}}\|_{1,\Omega} \tag{1.15}$$

and $\pi(\bar{\mathbf{v}}) = v$ (see [BBF13, Section 2.1.1]).

Since different parts of the boundary typically play different roles, we will now review the restriction properties of the discussed traces to boundary parts. This topic is also treated in [KO88, Chapter 5].

It is evident from the definiton of $H^{1/2}(\Gamma)$ that $H^{1/2}(\Gamma) \subset H^{1/2}(\Gamma_*)$ holds. Thus also the defintion of a local trace operator $\gamma_* : H^1(\Omega) \to H^{1/2}(\Gamma_*)$ as restriction of γ to Γ_* with the kernel $H^1_{\Gamma_*}(\Omega)$ is straightforward. We now consider the rest of the boundary $\Gamma_{**} := \mathrm{int}(\Gamma \setminus \Gamma_*)$ with the corresponding local trace γ_{**}. Since for $v \in H^1_{\Gamma_*}(\Omega)$ the global trace operator satisfies $\gamma(v) = 0$ on Γ_*, its restriction to Γ_{**} will always be extendable by zero outside of Γ_{**} suggesting $\gamma_{**} : H^1_{\Gamma_*}(\Omega) \to H^{1/2}_{00}(\Gamma_{**})$. Indeed it can be shown that γ_{**} is linear, continuous and surjective whenever the boundary of Γ_{**} is smooth (see [KO88, (5.31)]).

On the other hand, since $H^{-1/2}(\Gamma)$ is defined as a dual space, its elements cannot be restricted as simply and thus for $\mathbf{v} \in H(\mathrm{div}, \Omega)$ in general $\pi(\mathbf{v}) \notin H^{-1/2}(\Gamma_*)$ holds (see [BBF13, Section 2.5]). For this reason the definition of suitable stress-spaces in section 1.1.3 will be somewhat cumbersome. A general element v of $H^{-1/2}(\Gamma)$ can however be restricted to Γ_* if it is only paired with functions $u \in H^{1/2}_{00}(\Gamma_*)$. This is a consequence of the fact that every $u \in H^{1/2}_{00}(\Gamma_*)$ can by definition be extended by zero to $u_0 \in H^{-1/2}(\Gamma)$ such that

$$\|u_0\|_{1/2,\Gamma} \lesssim_{00} \|u\|_{1/2,\Gamma_*} \tag{1.16}$$

holds (see [KO88, Chapter 5]). The action of v on $u \in H^{1/2}_{00}(\Gamma_*)$ is simply given by

$${}_{00}\langle v, u \rangle_{\Gamma_*} = \langle v, u \rangle_{H^{-1/2}_{00}(\Gamma_*) \times H^{1/2}_{00}(\Gamma_*)} = \langle v, u_0 \rangle_{\Gamma}$$

Thus, it is important to make a distinction between $H^{-1/2}(\Gamma_*)$, which is the dual space of $H^{1/2}(\Gamma_*)$, and $H^{-1/2}_{00}(\Gamma_*)$, which is the dual space of $H^{1/2}_{00}(\Gamma_*)$. We point out that the extension of $v \in H^{-1/2}(\Gamma_*)$ by zero is always possible and we have the inverse inclusion $H^{-1/2}(\Gamma_*) \subset H^{-1/2}(\Gamma)$.

In the following we will simplify our notation by omittig the operators γ and π whenever it is clear that we are considering the traces of of functions in $H^1(\Omega)$ or $H(\mathrm{div}, \Omega)$ respectively.

1.1.3 Spaces related to Contact Problems

We will now define a number of subspaces and convex subsets of the Sobolev spaces introduced above which will be useful for the precise formulation of the problems we want to consider in this work. For this purpose we consider the following situation. Let

$\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz boundary which consists of three disjoint parts: $\partial\Omega = \Gamma = \overline{\Gamma_D \uplus \Gamma_N \uplus \Gamma_C}$, where we require Γ_D and Γ_N to be of positive measure. Let $\Gamma_R := \mathrm{int}(\overline{\Gamma_N \uplus \Gamma_C}) = \mathrm{int}(\Gamma \setminus \Gamma_D)$ and $\overline{\Gamma_C} \subset \Gamma_R$ hold. Furthermore let $\overline{\Gamma_C}$ be such, that the outward unit normal $\mathbf{n}$ on $\overline{\Gamma_C}$ is Lipschitz-continuous: $\mathbf{n} : \overline{\Gamma_C} \to \mathbb{R}^d \in \mathbf{C}^{0,1}(\overline{\Gamma_C})$. In this setting it is possible to separate the normal and tangential parts of a function $\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)$ as well as the normal and tangential parts of an element $\boldsymbol{\omega} \in \mathbf{H}^{-1/2}(\Gamma_C)$ (see [KO88, Theorems 5.4-5.11]). Indeed, $\mathbf{n} \in \mathbf{C}^{0,1}(\overline{\Gamma_C})$ and $\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)$ imply $(\mathbf{n}\cdot\mathbf{v}) \in H^{1/2}(\Gamma_C)$ (see [Gri85, Theorem 1.4.1.1]). Thus we can define for $\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)$:

$$\begin{aligned} v_n &:= \mathbf{n}\cdot\mathbf{v} \in H^{1/2}(\Gamma_C)\,, \\ \mathbf{v}_t &:= \mathbf{v} - v_n\mathbf{n} \in \mathbf{H}^{1/2}(\Gamma_C)\,, \end{aligned}$$

as well as for elements in the dual space $\boldsymbol{\omega} \in \mathbf{H}^{-1/2}(\Gamma_C)$:

$$\begin{aligned} \langle \omega_n, v\rangle_{\Gamma_C} &:= \langle \boldsymbol{\omega}, v\mathbf{n}\rangle_{\Gamma_C} \text{ for all } v \in H^{1/2}(\Gamma_C)\,, \\ \langle \boldsymbol{\omega}_t, \mathbf{v}\rangle_{\Gamma_C} &:= \langle \boldsymbol{\omega}, \mathbf{v}\rangle_{\Gamma_C} - \langle w_n, v_n\rangle_{\Gamma_C} \text{ for all } \mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)\,, \end{aligned}$$

and collect the tangential parts in the following spaces:

$$\begin{aligned} \mathbf{H}_T^{1/2}(\Gamma_C) &:= \{\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C) : v_n = 0\}\,, \\ \mathbf{H}_T^{-1/2}(\Gamma_C) &:= \{\boldsymbol{\omega} \in \mathbf{H}^{-1/2}(\Gamma_C) : \boldsymbol{\omega} = \boldsymbol{\omega}_t\}\,. \end{aligned}$$

Furthermore (using the ordering defined in [KO88, Chapter 5]), it is meaningful to define the following closed convex subsets of $H^{1/2}(\Gamma_C)$ and $H^{-1/2}(\Gamma_C)$ respectively:

$$P := \{v \in H^{1/2}(\Gamma_C) : v \geq 0\}\,, \tag{1.17}$$

$$P' := \{\omega \in H^{-1/2}(\Gamma_C) : \langle \omega, v\rangle_{\Gamma_C} \leq 0 \text{ for all } v \in P\}\,. \tag{1.18}$$

Since we also need to incorporate data into the spaces and sets related to contact problems we assume the following functions to be given: $\mathbf{f} \in \mathbf{L}_2(\Omega)$, $\mathbf{g}_D \in \mathbf{H}_{00}^{1/2}(\mathrm{int}(\overline{\Gamma_D \uplus \Gamma_N})) \subset \mathbf{H}^{1/2}(\Gamma)$, $\mathbf{g}_N \in \mathbf{L}_2(\Gamma_N) \subset \mathbf{H}_{00}^{-1/2}(\Gamma_R)$, $g_C \in H^{1/2}(\Gamma_C)$, $g_F \in P'$ and $\mu_F \geq 0$. We continue now with the definition of spaces and convex sets related to the primal variables:

$$\begin{aligned} \mathbf{U} &:= \mathbf{L}_2(\Omega)\,, \\ \mathbf{V} &:= \mathbf{H}^1_{\Gamma_D}(\Omega)\,, \\ \mathbf{V}_{\mathbf{g}_D} &:= \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v} = \mathbf{g}_D \text{ on } \Gamma\}\,, \\ \mathbf{K}_{g_C} &:= \{\mathbf{v} \in \mathbf{V} : g_C - v_n \in P\} = \{\mathbf{v} \in \mathbf{V} : g_C - v_n \geq 0 \text{ on } \Gamma_C\}\,, \\ \boldsymbol{\Theta} &:= \{\boldsymbol{\gamma} \in L_2(\Omega)^{d\times d} : \boldsymbol{\gamma} + \boldsymbol{\gamma}^{\mathsf{T}} = \mathbf{0}\}\,, \\ \mathbf{Q} &:= \mathbf{H}_T^{1/2}(\Gamma_C)\,. \end{aligned}$$

Before we can proceed to define the spaces and convex sets needed for the treatment of the dual variables, we have to take a closer look at the properties of the traces of

functions in $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}, \Omega)$. As discussed above in general the normal trace $\pi(\boldsymbol{\tau}) = \boldsymbol{\tau} \cdot \mathbf{n}$ can not be restricted to only part of the boundary, but we shall see that in our particular case it is possible to retreat to Γ_C. First we recall that for $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div}, \Omega)$ we have

$$\boldsymbol{\tau} \cdot \mathbf{n} \in \mathbf{H}^{-1/2}(\Gamma) \subset \mathbf{H}_{00}^{-1/2}(\Gamma_R) \tag{1.19}$$

with

$${}_{00}\|\boldsymbol{\tau} \cdot \mathbf{n}\|_{-1/2,\Gamma_R} \lesssim \|\boldsymbol{\tau} \cdot \mathbf{n}\|_{-1/2,\Gamma} \lesssim \|\boldsymbol{\tau}\|_{\mathrm{div},\Omega}\,. \tag{1.20}$$

For the duality pairing on Γ_R we simply get

$${}_{00}\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_R} = \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v}_0 \rangle_{\Gamma}\,, \tag{1.21}$$

where $\mathbf{v}_0$ is the extension by zero. Moreover if we prescribe the action of $\boldsymbol{\tau} \cdot \mathbf{n}$ on Γ_N we can localize the duality pairing even further: Assuming ${}_{00}\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_R} = (\mathbf{g}_N, \mathbf{v})_{L^2(\Gamma_N)}$ for all $\mathbf{v} \in \mathbf{H}_{00}^{1/2}(\Gamma_R)$ that satisfy $\mathrm{supp}(\mathbf{v}) \subset \Gamma_N$, we can define the action of $\boldsymbol{\tau} \cdot \mathbf{n}$ on Γ_C by

$$\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_C} := {}_{00}\langle \boldsymbol{\tau} \cdot \mathbf{n}, \tilde{\mathbf{v}} \rangle_{\Gamma_R} - (\mathbf{g}_N, \tilde{\mathbf{v}})_{L^2(\Gamma_N)}\,, \tag{1.22}$$

where $\tilde{\mathbf{v}}$ is a suitable extension of $\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)$ to $\mathbf{H}_{00}^{1/2}(\Gamma_R)$, that satisfies

$${}_{00}\|\tilde{\mathbf{v}}\|_{1/2,\Gamma_R} \lesssim \|\mathbf{v}\|_{1/2,\Gamma_C}\,. \tag{1.23}$$

Given the existence of such an extension (which holds whenever $\overline{\Gamma_C}$ is strictly included in Γ_R (see [DDE12, Chapter 4])), it is straightforward to see that $\langle \boldsymbol{\tau} \cdot \mathbf{n}, \cdot \rangle_{\Gamma_C}$ is well defined and bounded. Indeed for two extensions $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_2$ of $\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)$ we have $\tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_2 \equiv \mathbf{0}$ on Γ_C, thus $\mathrm{supp}(\tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_2) \subset \Gamma_N$ holds, implying

$${}_1\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_C} - {}_2\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_C} = {}_{00}\langle \boldsymbol{\tau} \cdot \mathbf{n}, \tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_2 \rangle_{\Gamma_R} - (\mathbf{g}_N, \tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_2)_{L^2(\Gamma_N)} = 0\,.$$

The boundedness is a consequence of (1.23) and (1.20):

$$\begin{aligned}
\|\boldsymbol{\tau} \cdot \mathbf{n}\|_{-1/2,\Gamma_C} &= \sup_{\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)} \frac{\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{v} \rangle_{\Gamma_C}}{\|\mathbf{v}\|_{1/2,\Gamma_C}} \\
&\lesssim \sup_{\mathbf{v} \in \mathbf{H}^{1/2}(\Gamma_C)} \frac{{}_{00}\langle \boldsymbol{\tau} \cdot \mathbf{n}, \tilde{\mathbf{v}} \rangle_{\Gamma_R} - (\mathbf{g}_N, \tilde{\mathbf{v}})_{L^2(\Gamma_N)}}{{}_{00}\|\tilde{\mathbf{v}}\|_{1/2,\Gamma_R}} \\
&\leq {}_{00}\|\boldsymbol{\tau} \cdot \mathbf{n}\|_{-1/2,\Gamma_R} + \|\mathbf{g}_N\|_{0,\Gamma_N}\,.
\end{aligned}$$

We conclude this section with the definition of spaces and convex sets related to the

dual variables:

$$
\begin{aligned}
\boldsymbol{\Sigma} &:= \mathbf{H}(\mathrm{div},\Omega)\,,\\
\boldsymbol{\Sigma}^{\mathbf{f}} &:= \{\boldsymbol{\tau}\in\boldsymbol{\Sigma} : \mathrm{div}\,\boldsymbol{\tau} = -\mathbf{f},\ \mathbf{as}\,\boldsymbol{\tau}=0\}\,,\\
\boldsymbol{\Sigma}_{\mathbf{g}_N} &:= \{\boldsymbol{\tau}\in\boldsymbol{\Sigma} : \langle\boldsymbol{\tau}\cdot\mathbf{n},\mathbf{v}\rangle_{\Gamma_R} = (\mathbf{g}_N,\mathbf{v})_{\Gamma_N}\ \forall\,\mathbf{v}\in\mathbf{H}^{1/2}_{00}(\Gamma_R)\}\,,\\
\boldsymbol{\Sigma}^F_{\mathbf{g}_N} &:= \{\boldsymbol{\tau}\in\boldsymbol{\Sigma} : \langle\boldsymbol{\tau}\cdot\mathbf{n},\mathbf{v}\rangle_{\Gamma_R} = (\mathbf{g}_N,\mathbf{v})_{\Gamma_N}\ \forall\,\mathbf{v}\in\mathbf{H}^{1/2}_{00}(\Gamma_R),\ \mathrm{supp}(\mathbf{v})\subset\Gamma_N\}\,,\\
\boldsymbol{\Sigma}^C_{\mathbf{g}_N} &:= \{\boldsymbol{\tau}\in\boldsymbol{\Sigma}^F_{\mathbf{g}_N} : \langle(\boldsymbol{\tau}\cdot\mathbf{n})_t,\mathbf{v}\rangle_{\Gamma_C} = 0\ \ \forall\,\mathbf{v}\in\mathbf{H}^{1/2}(\Gamma_C)\}\,,\\
\boldsymbol{\mathcal{K}}_{g_F} &:= \{\boldsymbol{\tau}\in\boldsymbol{\Sigma}^F_{\mathbf{0}} : (\boldsymbol{\tau}\cdot\mathbf{n})_n\in P',\ \langle(\boldsymbol{\tau}\cdot\mathbf{n})_t,\mathbf{r}\rangle_{\Gamma_C} + \langle\mu_F g_F,|\mathbf{r}|\rangle_{\Gamma_C}\le 0\ \forall\,\mathbf{r}\in\mathbf{Q}\}\,.
\end{aligned}
$$

Let $\boldsymbol{\sigma}_N\in\boldsymbol{\Sigma}_{\mathbf{g}_N}$, then we finally define

$$\boldsymbol{\Sigma}^* := (\boldsymbol{\sigma}_N+\boldsymbol{\Sigma}_{\mathbf{0}})\cap\boldsymbol{\Sigma}^{\mathbf{f}}\,, \tag{1.24}$$

$$\boldsymbol{\mathcal{K}}^*_{g_F} := (\boldsymbol{\sigma}_N+\boldsymbol{\mathcal{K}}_{g_F})\cap\boldsymbol{\Sigma}^{\mathbf{f}}\,, \tag{1.25}$$

and point out the following identities:

$$\boldsymbol{\Sigma}_{\mathbf{g}_N} = \boldsymbol{\sigma}_N+\boldsymbol{\Sigma}_{\mathbf{0}}\,, \tag{1.26}$$

$$\boldsymbol{\Sigma}^C_{\mathbf{g}_N} = \boldsymbol{\sigma}_N+\boldsymbol{\Sigma}^C_{\mathbf{0}}\,, \tag{1.27}$$

$$\boldsymbol{\Sigma}^F_{\mathbf{g}_N} = \boldsymbol{\sigma}_N+\boldsymbol{\Sigma}^F_{\mathbf{0}}\,, \tag{1.28}$$

$$\boldsymbol{\mathcal{K}}_{g_F}\subset\boldsymbol{\Sigma}^F_{\mathbf{0}}\,, \tag{1.29}$$

$$\boldsymbol{\mathcal{K}}_0\subset\boldsymbol{\Sigma}^C_{\mathbf{0}}\,. \tag{1.30}$$

Before we proceed we want to verify that both $\mathbf{K}_{g_C}$ and $\boldsymbol{\mathcal{K}}_{g_F}$ are indeed convex sets. The linearity of the trace-mapping directly yields

$$
\begin{aligned}
(\lambda\mathbf{v}+(1-\lambda)\mathbf{w})_n &= \gamma(\lambda\mathbf{v}+(1-\lambda)\mathbf{w})\cdot\mathbf{n}\\
&= \lambda\gamma(\mathbf{v})\cdot\mathbf{n}+(1-\lambda)\gamma(\mathbf{w})\cdot\mathbf{n} = \lambda v_n+(1-\lambda)w_n\le g_C
\end{aligned}
$$

for any $\mathbf{v},\mathbf{w}\in\mathbf{K}_{g_C}$. Similarly we make use of the fact that for $\boldsymbol{\sigma},\boldsymbol{\tau}\in\boldsymbol{\mathcal{K}}_{g_F}$ the restriction of the normal-trace is also linear. Consequently

$$
\begin{aligned}
\langle((\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n})_n,v\rangle_{\Gamma_C} &= \langle(\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n},v\mathbf{n}\rangle_{\Gamma_C} = (\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n},(v\tilde{\mathbf{n}})_0\rangle_{\Gamma}\\
&= \lambda\langle\boldsymbol{\sigma}\cdot\mathbf{n},(v\tilde{\mathbf{n}})_0\rangle_{\Gamma}+(1-\lambda)\langle\boldsymbol{\tau}\cdot\mathbf{n},(v\tilde{\mathbf{n}})_0\rangle_{\Gamma}\\
&= \lambda\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_n,v\rangle_{\Gamma}+(1-\lambda)\langle(\boldsymbol{\tau}\cdot\mathbf{n})_n,v\rangle_{\Gamma}\le 0\,.
\end{aligned}
$$

holds for all $v\in P$. In the same manner we also recover the constraint on the tangential component: For $\boldsymbol{\sigma},\boldsymbol{\tau}\in\boldsymbol{\mathcal{K}}_{g_F}$ and $\mathbf{r}\in\mathbf{Q}$ we have

$$
\begin{aligned}
\langle((\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n})_t,\mathbf{r}\rangle_{\Gamma_C} &= \langle(\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n},\mathbf{r}\rangle_{\Gamma_C}-\langle(\lambda\boldsymbol{\sigma}+(1-\lambda)\boldsymbol{\tau})\cdot\mathbf{n})_n,r_n\rangle_{\Gamma_C}\\
&= \lambda\langle\boldsymbol{\sigma}\cdot\mathbf{n},\tilde{\mathbf{r}}_0\rangle_{\Gamma}-\lambda\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_n,(\tilde{r_n})_0\rangle_{\Gamma}\\
&\quad+(1-\lambda)\langle\boldsymbol{\tau}\cdot\mathbf{n},\tilde{\mathbf{r}}_0\rangle_{\Gamma}-(1-\lambda)\langle(\boldsymbol{\tau}\cdot\mathbf{n})_n,(\tilde{r_n})_0\rangle_{\Gamma}\\
&= \lambda\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{r}\rangle_{\Gamma_C}+(1-\lambda)\langle(\boldsymbol{\tau}\cdot\mathbf{n})_t,\mathbf{r}\rangle_{\Gamma_C}\le-\langle\mu_F g_F,|\mathbf{r}|\rangle_{\Gamma_C}\,.
\end{aligned}
$$

1.2 Basics of Mixed FEM

In this section we will introduce some notation, review the definition of the finite element spaces used in this work and recapitulate some related results. Most of this can be found in [BBF13]. Let $\mathcal{T}_h = \{T_1, T_2, ...T_{n_t}\}$ be a shape-regular triangulation of Ω consisting of simplices (triangles in 2D, tetrahedra in 3D) with maximal diameter h. With $\mathcal{S}_h = \{S_1, S_2, ...S_{n_s}\}$ we denote the set of sides (edges in 2D, faces in 3D) and with $\mathcal{Z}_h = \{\mathbf{z}_1, \mathbf{z}_2, ...\mathbf{z}_{n_z}\}$ the set of vertices. We assume that $\mathcal{T}_h$ is compatible with the partition of Γ allowing the sound defintion of the set of sides of Γ_C:

$$S_{h,C} := \{S \in \mathcal{S}_h : S \subset \Gamma_C\} \tag{1.31}$$

1.2.1 FE Spaces

For $k \geq 0$, we denote by $\mathcal{P}_k(T)$ and $\mathcal{P}_k(S)$ the space of polynomials of maximal degree k on the simplex T or side S respectively. Using these local spaces we can define the following global spaces:

$$\mathcal{P}_k(\mathcal{T}_h) := \{v_h \in C(\Omega) : v_h \in \mathcal{P}_k(T) \ \forall\, T \in \mathcal{T}_h\}, \tag{1.32}$$
$$\mathcal{DP}_k(\mathcal{T}_h) := \{v_h \in L^2(\Omega) : v_h \in \mathcal{P}_k(T) \ \forall\, T \in \mathcal{T}_h\}, \tag{1.33}$$
$$\mathcal{P}_k(\mathcal{S}_{h,C}) := \{v_h \in C(\Gamma_C) : v_h \in \mathcal{P}_k(S) \ \forall\, S \in \mathcal{S}_{h,C}\}, \tag{1.34}$$
$$\mathcal{DP}_k(\mathcal{S}_{h,C}) := \{v_h \in L^2(\Gamma_C) : v_h \in \mathcal{P}_k(S) \ \forall\, S \in \mathcal{S}_{h,C}\}. \tag{1.35}$$

In analogy to (1.17) we also define

$$\mathcal{P}_k^+(\mathcal{S}_{h,C}) := \{v_h \in \mathcal{P}_k(\mathcal{S}_{h,C}) : v_h \geq 0\}, \tag{1.36}$$
$$\mathcal{DP}_k^+(\mathcal{S}_{h,C}) := \{v_h \in \mathcal{DP}_k(\mathcal{S}_{h,C}) : v_h \geq 0\}, \tag{1.37}$$
$$\mathcal{P}_k^-(\mathcal{S}_{h,C}) := -\mathcal{P}_k^+(\mathcal{S}_{h,C}), \tag{1.38}$$
$$\mathcal{DP}_k^-(\mathcal{S}_{h,C}) := -\mathcal{DP}_k^+(\mathcal{S}_{h,C}). \tag{1.39}$$

Furthermore we define the local space of Raviart-Thomas-functions of degree k:

$$\mathcal{RT}_k(T) := \{\mathbf{v}(\mathbf{x}) + \mathbf{x}\tilde{v}(\mathbf{x}) : \mathbf{v} \in \mathcal{P}_k(T)^d, \tilde{v} \in \mathcal{P}_k(T)\} \tag{1.40}$$

and recall the following properties of these functions.

Lemma 1.3. *For* $\mathbf{v} \in \mathcal{RT}_k(T)$

(i) $\operatorname{div} \mathbf{v} \in \mathcal{P}_k(T)$,

(ii) $(\mathbf{v} \cdot \mathbf{n})|_S \in \mathcal{P}_k(S) \quad \forall\, S \subset \partial T$,

(iii) $\operatorname{div} : \mathcal{RT}_k(T) \to \mathcal{P}_k(T)$ *is surjective,*

holds.

Proof. See [BBF13, Proposition 2.3.3]. □

The global Raviart-Thomas-spaces are then defined by

$$\mathcal{RT}_k(\mathcal{T}_h) := \{\mathbf{v}_h \in H(\mathrm{div},\Omega) : \mathbf{v}_h \in \mathcal{RT}_k(T)\ \forall\, T \in \mathcal{T}_h\}\,, \tag{1.41}$$

where the belonging to $H(\mathrm{div},\Omega)$ ensures the continuity of the normal component across internal sides. Since we will be especially concerned with the behaviour of $\mathcal{RT}_1$-functions on the Γ_C, we point out that (ii) in Lemma 1.3 implies

$$(\mathbf{v}\cdot\mathbf{n})|_{\Gamma_C} \in \mathcal{DP}_1(\mathcal{S}_{h,C}) \text{ for } \mathbf{v} \in \mathcal{RT}_1(\mathcal{T}_h)\,. \tag{1.42}$$

For the construction of an explicit basis of $\mathcal{RT}_1(\mathcal{T}_h)$ whose restriction to Γ_C yields the nodal basis of $\mathcal{DP}_1(\mathcal{S}_{h,C})$ we refer to [Kob17].

1.2.2 Interpolation estimates

Proposition 1.4. *Let $\mathcal{T}_h$ be a shape-regular triangulation. Then there exists an interpolation operator $\mathcal{I}_h^k : H^m(\Omega) \to \mathcal{P}_k(\mathcal{T}_h)$ such that*

$$\|v - \mathcal{I}_h^k v\|_s \le Ch^{m-s}|v|_m \quad 0 \le s \le m,\ 2 \le m \le k+1 \tag{1.43}$$

holds for a constant C independent of h.

Proof. See [BS08, Section 4.4]. □

Remark. $\mathcal{I}_h^k$ preserves the sign if $k = 1$:

$$v \ge 0 \text{ in } \Omega \ \Rightarrow\ \mathcal{I}_h^1 v \ge 0 \text{ in } \Omega\,. \tag{1.44}$$

Remark. It is possible to define an interpolation operator with comparable approximation properties also for $v \in H^1(\Omega)$ (see [BS08, Section 4.8],[Clé75]).

Definition 1.5. Let $V_h \subset L^2(\Omega)$ be a finite dimensional subspace of $L^2(\Omega)$, then the *L^2-orthogonal projection* $\pi_h : L^2(\Omega) \to V_h$ is defined by

$$\|v - \pi_h v\|_0 = \inf_{v_h \in V_h} \|v - v_h\|_0\,. \tag{1.45}$$

Remark. The L^2-orthogonal projection is characterized by

$$(v - \pi_h v, v_h) = 0 \quad \forall\, v_h \in V_h\,. \tag{1.46}$$

Proposition 1.6. *Let $\mathcal{T}_h$ be a shape-regular triangulation. Let $V_h \subset L^2(\Omega)$ be a finite dimensional subspace with $\mathcal{P}_k(\mathcal{T}_h) \subset V_h$. Then*

$$\|v - \pi_h v\|_s \le Ch^{m-s}|v|_m \quad 0 \le s \le m,\ 0 \le m \le k+1 \tag{1.47}$$

holds for a constant C independent of h.

Proof. See [DE12, Lemma 1.58]. □

Remark. While the parameters s and m in (1.43) and (1.47) are to be interpreted as integers in the first place, the techniques of real interpolation between Banach spaces (see for example [BS08, Chapter 14], [AF03, Chapter 7]) allow a generalization to estimates with non-integer Sobolev norms such as

$$\|v - \mathcal{I}_h^1 v\|_{1/2} \leq h^{\frac{3}{2}} \|v\|_2 \,. \tag{1.48}$$

Moreover, using duality arguments it is even possible to obtain estimates involving negative Sobolev norms such as the following:

$$\begin{aligned} \|v - \pi_h v\|_{-1/2} &= \sup_{w \in H^{1/2}(\Omega)} \frac{\langle v - \pi_h v, w \rangle_\Omega}{\|w\|_{1/2}} \\ &= \sup_{w \in H^{1/2}(\Omega)} \frac{(v - \pi_h v, w - \pi_h w)}{\|w\|_{1/2}} \\ &\leq \sup_{w \in H^{1/2}(\Omega)} \frac{\|v - \pi_h v\|_0 \|w - \pi_h w\|_0}{\|w\|_{1/2}} \\ &\leq \sup_{w \in H^{1/2}(\Omega)} \frac{C_1 h^2 \|v\|_2 C_2 h^{\frac{1}{2}} \|w\|_{1/2}}{\|w\|_{1/2}} = C h^{\frac{5}{2}} \|v\|_2 \end{aligned} \tag{1.49}$$

Proposition 1.7. *Let $\mathcal{T}_h$ be a shape-regular triangulation. Then there exists an interpolation operator $\mathcal{R}_h^k : H(\mathrm{div}, \Omega) \cap L^r(\Omega) \to \mathcal{RT}_k(\mathcal{T}_h)$ (with $r > 2$ fixed) such that*

$$\begin{aligned} \|\mathbf{v} - \mathcal{R}_h^k \mathbf{v}\|_0 &\leq C h^m |\mathbf{v}|_m \quad 1 \leq m \leq k+1 \\ \|\mathrm{div}(\mathbf{v} - \mathcal{R}_h^k \mathbf{v})\|_0 &\leq C h^s |\mathrm{div}\, \mathbf{v}|_m \quad 0 \leq s \leq k+1 \end{aligned} \tag{1.50}$$

holds for a constant C independent of h. Furthermore

$$\mathrm{div}(\mathcal{R}_h^k \mathbf{v}) = \pi_h^k(\mathrm{div}, \mathbf{v}) \tag{1.51}$$

holds, with π_h^k being the L^2-orthogonal projection onto $\mathcal{DP}_k(\mathcal{T}_h)$.

Proof. See [BBF13, Section 2.5.1]. □

The interpolation operator of Proposition 1.7 is defined by

$$\int_S (\mathbf{v} - \mathcal{R}_h^k) \cdot \mathbf{n}\, p_k \, ds = 0 \quad \forall\, p_k \in \mathcal{P}_k(S),\, S \subset \partial T \,, \tag{1.52}$$

$$\int_T (\mathbf{v} - \mathcal{R}_h^k)\, \mathbf{p}_{k-1} \, dx = 0 \quad \forall\, \mathbf{p}_{k-1} \in \mathcal{P}_{k-1}(T)^d \tag{1.53}$$

(see [BBF13, Section 2.5.1]).

1.3 Miscellaneous

Useful Inequalities

In this subsection we will recapitulate a number of inequalities that are tightly linked to the analysis of linear elasticity and thus are also relevant for the treatment of contact problems. We decided to group them in this chapter for easier reference, even though some of the used symbols will not be introduced before the subsequent chapter. The setting is as described at the beginning of section 1.1.3.

First we recall the well known Korn's inequality.

Lemma 1.8. *There exists a constant C_K such that for every $\mathbf{v} \in \mathbf{V} = \mathbf{H}^1_{\Gamma_D}(\Omega)$ we have*

$$\|\mathbf{v}\|_1 \leq C_K \|\boldsymbol{\varepsilon}(\mathbf{v})\|_0 \tag{1.54}$$

holds.

Proof. See [KO88, Chapter 5]. □

Next we review an inequality for matrices with rows in $H(\mathrm{div})$ and a useful consequence thereof for elasticity problems.

Lemma 1.9. *There exists a constant $\tilde{C_D}$ such that for every $\boldsymbol{\tau} \in \boldsymbol{\Sigma_0} = \mathbf{H}_{\Gamma_N}(\mathrm{div})(\Omega)$ we have*

$$\|\mathrm{tr}\,\boldsymbol{\tau}\|_0 \leq \tilde{C}_D(\|\mathbf{dev}\,\boldsymbol{\tau}\|_0 + \|\mathrm{div}\,\boldsymbol{\tau}\|_0)\,. \tag{1.55}$$

Proof. See [BBF13, Section 9.1.1]. □

Corollary 1.10. *There exists a constant C_D such that for every $\boldsymbol{\tau} \in \boldsymbol{\Sigma_0}$ with $\mathrm{div}\,\boldsymbol{\tau} = \mathbf{0}$ we have*

$$\|\mathrm{tr}\,\boldsymbol{\tau}\|_0 \leq C_D \|\mathcal{A}\boldsymbol{\tau}\|_0\,. \tag{1.56}$$

Proof. See [CS03, Section 5]. □

Finally we will often use the following elementary special case of Young's inequality valid for $a, b \in \mathbb{R}$ and $\delta > 0$:

$$ab \leq \frac{\delta}{2}a^2 + \frac{1}{2\delta}b^2\,. \tag{1.57}$$

Minimization theorems

We quote here the generalized Weierstrass minimization theorem (which can be found for example in [OK80]), followed by a corollary that is tailored to most of the problems which we will consider in this work.

Theorem 1.11. *Let V be a reflexive Banach space and K a non-empty closed convex subset of V. Let $f : K \to \mathbb{R}$ be a functional defined on K which is weakly lower semicontinuous, i.e. :*

If $\{u_m\} \in K$ converges weakly to $u \in K$, then

$$\liminf_{m\to\infty} f(u_m) \geq f(u)\,. \tag{1.58}$$

Then f is bounded below on K and attains its minimum value on K whenever either of the following conditions hold:

(i) K is bounded, or

(ii) f is coercive, i.e.

$$f(v) \to +\infty \text{ as } \|v\| \to \infty\ ,. \tag{1.59}$$

Proof. See [OK80, Theorem 1-2.2]. □

Corollary 1.12. *Let $K \subset V$ be a non-empty, convex and closed subset of a Hilbert space V. Let f be quadratic, i.e. of the form*

$$f(v) = a(v,v) + b(v) \tag{1.60}$$

with a continuous symmetric bilinear form $a(\cdot,\cdot)$ on $V \times V$ and a continuous linear form b on V. Let furthermore $a(\cdot,\cdot)$ be elliptic on K i.e.

$$\|v\|^2 \lesssim a(v,v) \quad \forall\, v \in K \tag{1.61}$$

then

$$\min_{x\in K} f(x) \tag{1.62}$$

admits an unique solution.

Proof. As a Hilbert space V is a reflexive Banach space. The ellipticity of $a(\cdot,\cdot)$ implies the coercivity and strict convexity of f. Furthermore the continuity of $a(\cdot,\cdot)$ and b implies the continuity of f which together with the convexity of f implies the weak lower semicontinuity of f. Thus Theorem 1.11 guarantees the existence of a solution of (1.62) and the strict convexity of f implies its uniqueness. □

Chapter 2

A stress-based approach to Linear Elasticity

2.1 Linear Elasticity

The starting point of this thesis is the well researched problem of linear elasticity. We will include only a short introduction in order to set the stage for our study of contact problems, and refer the reader to [Cia88, Bra13, KO88] for a more extensive treatment of this topic. We are interested in the deformation and resultant internal stresses of an elastic solid subject to a given loading. In the theory of *Linear* Elasticity the arising strains are assumed to be small, justifying linear relationships between displacements, strains and stresses.

To fix ideas, let the reference configuration of an elastic solid be represented by a bounded subset of Euclidean space $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ with boundary $\Gamma := \partial\Omega$. Let it be subject to the body force $\mathbf{f} : \Omega \to \mathbb{R}^d$ acting throughout its volume and the surface force $\mathbf{g}_N : \Gamma_N \to \mathbb{R}^d$ acting on a part of the boundary $\Gamma_N \subset \Gamma$. The rest of the boundary $\Gamma_D := \Gamma \setminus \overline{\Gamma_N}$ is clamped (or a displacement $\mathbf{g}_D$ is prescribed thereon). The quantities of interest are the resulting displacement $\mathbf{u} : \Omega \to \mathbb{R}^d$ and stress fields $\boldsymbol{\sigma} : \Omega \to \mathbb{R}^{d\times d}$, which are characterized by the following equations.

The kinematic relationship between displacement and strains is expressed by the defintion of the linearized strain tensor $\boldsymbol{\varepsilon}$ as the symmetric part of the gradient of the displacements:

$$\varepsilon(\mathbf{u}) := \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^{\mathsf{T}}\right). \tag{2.1}$$

The conservation of linear and angular momentum is expressed by Cauchy's balance laws for the stress tensor:

$$\operatorname{div}\boldsymbol{\sigma} = -\mathbf{f} \qquad \text{in } \Omega, \tag{2.2a}$$

$$\mathbf{as}\,\boldsymbol{\sigma} = \mathbf{0} \qquad \text{in } \Omega, \tag{2.2b}$$

where $\mathbf{as}\,\boldsymbol{\tau} := \frac{1}{2}(\boldsymbol{\tau} - \boldsymbol{\tau}^{\mathsf{T}})$ denotes the antisymmetric part of a matrix. The constitutive relation of an isotropic and homogenous material is expressed by Hooke's law featuring

the material dependent Lamé parameters $\mu, \lambda > 0$:

$$\boldsymbol{\sigma} = \mathcal{C}\boldsymbol{\varepsilon} := 2\mu\boldsymbol{\varepsilon} + \lambda(\operatorname{tr}\boldsymbol{\varepsilon})\mathbf{I}\,. \tag{2.3}$$

The Lamé parameters quantify the materials resistance to shear stress (μ) and volume change (λ). For $\lambda < \infty$ the operator $\mathcal{C}$ (usually called *stiffness tensor*) admits the inverse

$$\boldsymbol{\varepsilon} = \mathcal{C}^{-1}\boldsymbol{\sigma} := \frac{1}{2\mu}\left(\boldsymbol{\sigma} - \frac{\lambda}{d\lambda + 2\mu}(\operatorname{tr}\boldsymbol{\sigma})\mathbf{I}\right)$$

which has the advantage of remaining bounded also for incompressible materials ($\lambda = \infty$). This allows the formulation of the constitutive law in terms of stresses in the following way:

$$\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\varepsilon}, \qquad \text{with} \quad \mathcal{A} := \begin{cases} \mathcal{C}^{-1}, & \lambda < \infty \\ \dfrac{1}{2\mu}\mathbf{dev}, & \lambda = \infty \end{cases} \tag{2.4}$$

where **dev** denotes the deviatoric part of a matrix. $\mathcal{A}$ is usually called *compliance tensor*. Note that (2.4) already implies (2.2b). However it will be useful, even necessary, to retain the symmetry constraint explicitly during the following manipulations. Finally the boundary conditions complete the system of equations of linear elasticity:

$$\boldsymbol{\sigma}\cdot\mathbf{n} = \mathbf{g}_N \qquad \text{on } \Gamma_N, \tag{2.5a}$$

$$\mathbf{u} = \mathbf{g}_D \qquad \text{on } \Gamma_D. \tag{2.5b}$$

We listed them last, because they, or, more accurately, their modification, will be the point of interest during the subsequent chapters dealing with contact problems. The equations (2.1) - (2.4), however, will remain the same throughout, forming the basis of our considerations. Before we discuss some generalized formulations of the problem of linear elasticity we summarize the *classical problem of linear elasticity*:

Find sufficiently smooth $\mathbf{u}$ and $\boldsymbol{\sigma}$ such that

$$\begin{aligned} \operatorname{div}\boldsymbol{\sigma} &= -\mathbf{f} && \text{in } \Omega, \\ \mathbf{as}\,\boldsymbol{\sigma} &= \mathbf{0} && \text{in } \Omega, \\ \mathcal{A}\boldsymbol{\sigma} &= \boldsymbol{\varepsilon}(\mathbf{u}) && \text{in } \Omega, \\ \boldsymbol{\sigma}\cdot\mathbf{n} &= \mathbf{g}_N && \text{on } \Gamma_N, \\ \mathbf{u} &= \mathbf{g}_D && \text{on } \Gamma_D\,. \end{aligned} \tag{2.6}$$

In the classical formulation the equations are to be understood in a pointwise sense, while in the following sections equations in Ω are to be understood in an "almost everywhere-" or L^2-sense and equations on Γ or its parts are to be understood in the sense of traces.

2.2 Variational formulations

In this section we will discuss a number of variational formulations of the problem of linear elasticity, their solvability and their connection to each other. First we will touch

briefly on the most common approach omitting the details of its derivation, since very similar steps will be taken in the derivation of the less common stress-based formulation on which we will base our finite element method.

Displacement formulation

The standard way to derive a variational formulation of (2.6) is to multiply (2.2a) with a test function $\mathbf{v}$, integrate over Ω and apply partial integration arriving at

$$\int_\Omega \boldsymbol{\sigma} : \boldsymbol{\nabla}\mathbf{v}\, dx = \int_\Omega \mathbf{f} : \mathbf{v}\, dx + \int_\Gamma (\boldsymbol{\sigma}\cdot\mathbf{n})\cdot\mathbf{v}\, ds\,,$$

where ":" denotes the Frobenius inner product. Inserting the boundary condition (2.5a), substituting the combined equations (2.1) - (2.3) and choosing appropriate function spaces we obtain the following problem:

Find $\mathbf{u} \in \mathbf{V}_{\mathbf{g}_D}$ such that

$$(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}))_\Omega = (\mathbf{f},\mathbf{v})_\Omega + (\mathbf{g}_N,\mathbf{v})_{\Gamma_N} \quad \forall\, \mathbf{v} \in \mathbf{V}\,. \tag{2.7}$$

(2.7) can also be interpreted as first order optimality conditions of an optimization problem that seeks to minimize the internal energy dependent on the displacements and which we will subsequently refer to as the *primal problem of linear elasticity*:

Find $\mathbf{u} \in \mathbf{V}_{\mathbf{g}_D}$ such that

$$\mathcal{J}_{pr}(\mathbf{u}) = \min_{\mathbf{v}\in\mathbf{V}_{\mathbf{g}_D}} \mathcal{J}_{pr}(\mathbf{v}) := \frac{1}{2}(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{v}), \boldsymbol{\varepsilon}(\mathbf{v}))_\Omega - (\mathbf{f},\mathbf{v})_\Omega - (\mathbf{g}_N,\mathbf{v})_{\Gamma_N}\,. \tag{2.8}$$

Remark. In the fully incompressible case ($\lambda = \infty$) the primal problem (2.7) is no longer well defined since $\mathcal{C}$ is unbounded. In this case it is necessary to introduce an auxiliary pressure variable to obtain a well posed primal formulation [Bra13, Chapter VI]:

Find $\mathbf{u} \in \mathbf{V}_{\mathbf{g}_D}$ and $p \in L^2(\Omega)$ such that

$$\begin{aligned} 2\mu(\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}))_\Omega + (p, \operatorname{div}\mathbf{v})_\Omega &= (\mathbf{f},\mathbf{v})_\Omega + (\mathbf{g}_N,\mathbf{v})_{\Gamma_N} \quad \forall\, \mathbf{v} \in \mathbf{V}\,, \\ (\operatorname{div}\mathbf{u}, q)_\Omega &= 0\,. \end{aligned} \tag{2.9}$$

Stress-based formulation

An alternative approach to obtain a variational formulation of the problem of linear elasticity starts with the weak form of the constitutive equation in terms of stresses (2.4):

$$\int_\Omega (\mathcal{A}\boldsymbol{\sigma}) : \boldsymbol{\tau}\, dx = \int_\Omega \boldsymbol{\varepsilon} : \boldsymbol{\tau}\, dx \quad \forall\, \boldsymbol{\tau} \in C^\infty(\Omega)^{d\times d} \text{ s.t. } \boldsymbol{\tau}\cdot\mathbf{n} = \mathbf{0} \text{ on } \Gamma_N\,.$$

Inserting (2.1) as $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u}) = \boldsymbol{\nabla}\mathbf{u} - \mathbf{as}(\boldsymbol{\nabla}\mathbf{u})$ and applying partial integration we obtain

$$\begin{aligned} \int_\Omega (\mathcal{A}\boldsymbol{\sigma}) : \boldsymbol{\tau}\, dx &= \int_\Omega (\boldsymbol{\nabla}\mathbf{u} - \mathbf{as}\,(\boldsymbol{\nabla}\mathbf{u})) : \boldsymbol{\tau}\, dx \\ &= -\int_\Omega \mathbf{u}\cdot(\operatorname{div}\boldsymbol{\tau})\, dx + \int_\Gamma \mathbf{u}\cdot(\boldsymbol{\tau}\cdot\mathbf{n})\, ds - \int_\Omega \mathbf{as}\,(\boldsymbol{\nabla}\mathbf{u}) : \boldsymbol{\tau}\, dx\,. \end{aligned}$$

Introducing the new variable $\boldsymbol{\theta} := \mathbf{as}(\boldsymbol{\nabla}\mathbf{u})$ (which is sometimes referred to as "rotations" since local rotations are linked to the antisymmetric part of the displacement gradient) and inserting the boundary condition (2.5b) yields

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau})_\Omega + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau})_\Omega + (\boldsymbol{\theta}, \mathbf{as}\,\boldsymbol{\tau})_\Omega = \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma .$$

Finally we add the weak forms of the balance laws (2.2a)-(2.2b) and extend the ansatz and test spaces to include all functions for which the used expressions are well defined. The Neumann boundary condition is integrated in the ansatz space. This yields the so called generalized Hellinger-Reissner principle:

Find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\mathbf{g}_N}$, $\mathbf{u} \in \mathbf{U}$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ such that

$$\begin{aligned} (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\,\boldsymbol{\tau}) &= \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma \qquad && \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}, \\ (\operatorname{div}\boldsymbol{\sigma}, \mathbf{v}) &= (-\mathbf{f}, \mathbf{v}) && \forall\, \mathbf{v} \in \mathbf{U}, \\ (\mathbf{as}\,\boldsymbol{\sigma}, \boldsymbol{\gamma}) &= 0 && \forall\, \boldsymbol{\gamma} \in \boldsymbol{\Theta}. \end{aligned} \tag{2.10}$$

For practical reasons (such as implementation) it is often useful to reduce (2.10) to the case with homogeneous boundary conditions on Γ_N. For this purpose, let any $\boldsymbol{\sigma}_N \in \boldsymbol{\Sigma}_{\mathbf{g}_N}$ be given, then (2.10) is equivalent to finding $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_N + \hat{\boldsymbol{\sigma}}) \in (\boldsymbol{\sigma}_N + \boldsymbol{\Sigma}_{\mathbf{0}})$, $\mathbf{u} \in \mathbf{U}$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ such that

$$\begin{aligned} (\mathcal{A}\hat{\boldsymbol{\sigma}}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\,\boldsymbol{\tau}) &= \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma - (\mathcal{A}\boldsymbol{\sigma}_N, \boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}, \\ (\operatorname{div}\hat{\boldsymbol{\sigma}}, \mathbf{v}) &= -(\mathbf{f} + \operatorname{div}\boldsymbol{\sigma}_N, \mathbf{v}) && \forall\, \mathbf{v} \in \mathbf{U}, \\ (\mathbf{as}\,\hat{\boldsymbol{\sigma}}, \boldsymbol{\gamma}) &= -(\mathbf{as}\,\boldsymbol{\sigma}_N, \boldsymbol{\gamma}) && \forall\, \boldsymbol{\gamma} \in \boldsymbol{\Theta}. \end{aligned} \tag{2.11}$$

To simplify notation we will mostly stick to the notation with non-homogeneous boundary conditions (2.10) and ask the reader to keep in mind the straightforward interpretation of $\boldsymbol{\Sigma}_{\mathbf{g}_N} = (\boldsymbol{\sigma}_N + \boldsymbol{\Sigma}_{\mathbf{0}})$ turning it into (2.11).

Of course one could include the balance laws explicitly in the ansatz space in order to avoid a mixed formulation: Find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^*$ such that

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau})_\Omega = \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}^{\mathbf{0}} \cap \boldsymbol{\Sigma}_{\mathbf{0}} . \tag{2.12}$$

However, since this approach is hardly practicable for finite element discretizations, we will continue to work with (2.10).

Just like in the primal case, (2.12) can be interpreted as optimality condition of minimizing the internal energy, but this time in terms of stresses. We will refer to this as the *dual problem of linear elasticity*:

Find $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^*$ such that

$$\mathcal{J}_{du}(\boldsymbol{\sigma}) = \min_{\boldsymbol{\tau}\in\boldsymbol{\Sigma}^*} \mathcal{J}_{du}(\boldsymbol{\tau}) := \frac{1}{2}(\mathcal{A}\boldsymbol{\tau}, \boldsymbol{\tau})_\Omega - \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma + (\mathbf{g}_N, \mathbf{g}_D)_{\Gamma_N} . \tag{2.13}$$

Thus the variables $\mathbf{u}$ and $\boldsymbol{\theta}$ can be interpreted as the Lagrange multipliers defined by $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\theta})$ being the saddle point of the Lagrangian

$$\mathcal{L}(\boldsymbol{\tau}, \mathbf{v}, \boldsymbol{\gamma}) := \mathcal{J}_{du}(\boldsymbol{\tau}) + (\mathbf{v}, \operatorname{div}\boldsymbol{\tau} + \mathbf{f}) + (\boldsymbol{\gamma}, \mathbf{as}\,\boldsymbol{\tau}) . \tag{2.14}$$

The unique existence of the Lagrange multipliers for a given solution $\boldsymbol{\sigma}$ is evident from using the Riesz-representation of $(\mathcal{A}\boldsymbol{\sigma}, \cdot)_\Omega - \langle \cdot, \mathbf{g}_D\rangle_\Gamma$ in $H(\operatorname{div}, \Omega)$.

Remark. The last term in the definition of $\mathcal{J}_{du}(\boldsymbol{\tau})$ in (2.13) obviously can (and subsequently will) be omitted since it is independent of $\boldsymbol{\tau}$.

2.2.1 Duality

The relationship between the dual and the primal problem of linear elasticity can be expressed as follows:

Theorem 2.1. *Let $\lambda < \infty$.*

(i) Let $\mathbf{u}$ be the solution of (2.7), then $\boldsymbol{\sigma}(\mathbf{u}) := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u})$, $\mathbf{u}$ and $\boldsymbol{\theta}(\mathbf{u}) := \mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}$ solve (2.10).

(ii) Let conversely $\boldsymbol{\sigma}$, $\mathbf{u}$ and $\boldsymbol{\theta}$ be the solution of (2.10), then $\mathbf{u}$ is in $H^1(\Omega)$ and solves (2.7).

Proof. (i) Let $\mathbf{u}$ be the solution of (2.7). Then $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}) \in L^2(\Omega)^{d\times d}$ and $\mathbf{as}\,\boldsymbol{\sigma} = \mathbf{0}$. Testing with $\mathbf{v} = \boldsymbol{\phi}$ with $\boldsymbol{\phi} \in C_c^\infty(\Omega)^d$ in (2.7) we obtain $\operatorname{div}\boldsymbol{\sigma} = -\mathbf{f}$ implying $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^{\mathbf{f}}$ and

$$\langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{v}\rangle_\Gamma = (\mathbf{g}_N, \mathbf{v})_{\Gamma_N} \quad \forall\, \mathbf{v} \in \mathbf{V}\,,$$

which is equivalent to $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}_{\mathbf{g}_N}$. Thus $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^*$ and what remains to be shown is the first equation in (2.10). Setting $\boldsymbol{\theta} := \mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}$, inserting all variables into the equation and applying partial integration leads to

$$\begin{aligned}(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\,\boldsymbol{\tau}) &= (\mathcal{A}\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}, \mathbf{as}\,\boldsymbol{\tau})\\ &= (\boldsymbol{\varepsilon}(\mathbf{u}) - \boldsymbol{\nabla}\mathbf{u} + \mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}, \boldsymbol{\tau}) + \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{u}\rangle_\Gamma\\ &= \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma\end{aligned}$$

for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}$, where we used $\mathbf{u} \in \mathbf{V}_{\mathbf{g}_D}$ and $\mathcal{A} = \mathcal{C}^{-1}$.

(ii) Let now $\boldsymbol{\sigma}$, $\mathbf{u}$ and $\boldsymbol{\theta}$ solve (2.10). By testing only with $\boldsymbol{\tau} \in C_c^\infty(\Omega)^{d\times d}$ in the first equation in (2.10) we obtain

$$(\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) = -(\mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}, \boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in C_c^\infty(\Omega)^{d\times d} \tag{2.15}$$

implying $\mathbf{u} \in H^1(\Omega)$ with $\boldsymbol{\nabla}\mathbf{u} = \mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}$. Consequently partial integration yields

$$\langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{u}\rangle_\Gamma = \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}\,. \tag{2.16}$$

implying $\mathbf{u} \in \mathbf{V}_{g_D}$. Finally observing $\mathbf{V} \subset \mathbf{U}$ and $\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}) = \boldsymbol{\sigma}$ (which is a consequence of $\mathbf{as}\,\boldsymbol{\sigma} = \mathbf{0}$), partial integration of the second equation in (2.10) yields (2.7). □

Remark. This duality result stays true also in the incompressible limit ($\lambda = \infty$), with the adaptation that the stress is defined as $\boldsymbol{\sigma}(\mathbf{u}, p) := 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) + \mathbf{I}p$, conserving the equality $\mathcal{A}\boldsymbol{\sigma}(\mathbf{u}, p) = \boldsymbol{\varepsilon}(\mathbf{u})$.

Proof. (i) can be shown analogous to the proof in the theorem. For *(ii)* we need to define the pressure as $p := \frac{1}{d}\mathrm{tr}(\boldsymbol{\sigma})$. This yields $p \in L^2(\Omega)$ and following the same steps as for the compressible case we obtain the first equation in (2.9). The second equation is verified in the following way:

$$\operatorname{div} \mathbf{u} = \mathrm{tr}(\boldsymbol{\nabla}\mathbf{u}) = \mathrm{tr}(\mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}) = \mathrm{tr}(\mathcal{A}\boldsymbol{\sigma}) = \frac{1}{2\mu}\mathrm{tr}(\mathbf{dev}\,\boldsymbol{\sigma}) = 0\,.$$

□

2.2.2 Existence and Uniqueness

Since $\mathcal{J}_{du}(\boldsymbol{\tau})$ is quadratic (in the sense of Theorem 1.12) the well-posedness of (2.13) rests on the nonemptyness of the closed affine subspace $\boldsymbol{\Sigma}^*$ and the coercivity of the bilinearform $(\mathcal{A}\cdot,\cdot)$ thereon. The former is usually expressed in the form of an inf-sup-condition:

$$\inf_{\substack{\mathbf{v}\in\mathbf{U}\\ \boldsymbol{\gamma}\in\boldsymbol{\Theta}}} \sup_{\boldsymbol{\tau}\in\boldsymbol{\Sigma_0}} \frac{(\operatorname{div}\boldsymbol{\tau},\mathbf{v}) + (\mathbf{as}\,\boldsymbol{\tau},\boldsymbol{\gamma})}{\|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}(\|\mathbf{v}\|_{\mathbf{U}} + \|\boldsymbol{\gamma}\|_{\boldsymbol{\Theta}})} \geq \beta > 0\,. \tag{2.17}$$

Both the validity of (2.17) and coercivity, that is

$$\|\boldsymbol{\tau}\|^2_{\boldsymbol{\Sigma}} \lesssim (\mathcal{A}\boldsymbol{\tau},\boldsymbol{\tau}) \quad \forall\,\boldsymbol{\tau} \in \{\boldsymbol{\tau} \in \boldsymbol{\Sigma_0} : \operatorname{div}\boldsymbol{\tau} = \mathbf{0}\} \tag{2.18}$$

with a constant independent of λ, are discussed and verified in [BBF13] and [BBF09]. (2.18) implies the strict convexity of $\mathcal{J}_{du}(\boldsymbol{\tau})$ on $\boldsymbol{\Sigma}^*$, hence Theorem 1.12 applies, yielding the existence of a unique solution of (2.13). Consequently (2.10) are solvable, sufficient (and necessary) optimality conditions. Theorem 2.1 makes clear that the primal problem is well posed if and only if the dual problem is well posed. We summarize this in the following theorem:

Theorem 2.2. *Both the primal problem of linear elasticity (2.8) or equivalently (2.7) and the dual problem of linear elasticity (2.13) or equivalently (2.10) admit a unique solution.*

Finally we point out that it is evident from our derivation, that every solution of (2.6) also solves (2.10) with $\boldsymbol{\theta} = \mathbf{as}(\boldsymbol{\nabla}\mathbf{u})$. On the other hand one can easily verify that a sufficiently smooth solution of (2.10) also solves the classical problem (2.6).

2.3 Finite Element Discretization

In this section we will present a short overview on some possible choices for the finite element discretization of (2.10) and briefly discuss their advantages and disadvantages. First we write down the general discretized problem:

Find $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_{\mathbf{g}_N,h}$, $\mathbf{u}_h \in \mathbf{U}_h$ and $\boldsymbol{\theta}_h \in \boldsymbol{\Theta}_h$ such that

$$\begin{aligned}
(\mathcal{A}\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h) + (\mathbf{u}_h,\operatorname{div}\boldsymbol{\tau}_h) + (\boldsymbol{\theta}_h,\mathbf{as}\,\boldsymbol{\tau}_h) &= \langle \boldsymbol{\tau}_h\cdot\mathbf{n},\mathbf{g}_D\rangle_\Gamma \quad && \forall\,\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_{\mathbf{0},h},\\
(\operatorname{div}\boldsymbol{\sigma}_h,\mathbf{v}_h) &= (-\mathbf{f},\mathbf{v}_h) && \forall\,\mathbf{v}_h \in \mathbf{U}_h,\\
(\mathbf{as}\,\boldsymbol{\sigma}_h,\boldsymbol{\gamma}_h) &= 0 && \forall\,\boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h.
\end{aligned} \tag{2.19}$$

(2.19) can be interpreted as the optimality conditions of minimizing $\mathcal{J}_{du}$ on the discrete admissible set

$$\begin{aligned}\boldsymbol{\Sigma}_h^* := \{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_{\mathbf{g}_N,h} : (\operatorname{div}\boldsymbol{\sigma}_h, \mathbf{v}_h) = (-\mathbf{f}, \mathbf{v}_h) \quad &\forall\, \mathbf{v}_h \in \mathbf{U}_h,\\ (\mathbf{as}\,\boldsymbol{\sigma}_h, \boldsymbol{\gamma}_h) = 0 \quad &\forall\, \boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h\}\end{aligned} \tag{2.20}$$

The main challenge of this formulation is the sound treatment of the symmetry constraint while retaining enough degrees of freedom to satisfy the conservation of linear momentum. Traditionally a strong symmetry constraint is integrated into the ansatz space for the stress approximations $\boldsymbol{\Sigma}_h$ such that the third equation in (2.19) can be omitted. The methods discussed in [AAW08, AW02] and [GG11] are examples of this approach. However the Arnold-Winther-Element is rather complicated and computationally expensive, since even in the lowest-order case polynomials of degree three (in 2D) and four (in 3D) are needed. Gopalakrishnan and Guzman are able to reduce the necessary polynomial degree to two but they have to pay for this with suboptimal convergence rates and non-$H(\operatorname{div})$-conforming elements (see also [KS17]). Thus we turn to the so called *reduced* or *weak* symmetry approach that allows non symmetric stress approximations while controlling the antisymmetric part with an appropriate choice of $\boldsymbol{\Theta}_h$. The first methods following this approach were analyzed in [ABD84] and [AT79]. In [BBF09] a number of combinations of $(\boldsymbol{\Sigma}_h, \mathbf{U}_h, \boldsymbol{\Theta}_h)$ are discussed, that satisfy a discrete inf-sup-condition analogous to (2.17):

$$\inf_{\substack{\mathbf{v}_h\in\mathbf{U}_h\\ \boldsymbol{\gamma}_h\in\boldsymbol{\Theta}_h}} \sup_{\boldsymbol{\tau}_h\in\boldsymbol{\Sigma}_{\mathbf{0},h}} \frac{(\operatorname{div}\boldsymbol{\tau}_h, \mathbf{v}_h) + (\mathbf{as}\,\boldsymbol{\tau}_h, \boldsymbol{\gamma}_h)}{\|\boldsymbol{\tau}_h\|_{\boldsymbol{\Sigma}}(\|\mathbf{v}_h\|_{\mathbf{U}} + \|\boldsymbol{\gamma}_h\|_{\boldsymbol{\Theta}})} \geq \beta > 0\,. \tag{2.21}$$

For the generalization of (2.19) to contact problems we chose to focus on the following inf-sup stable combination presented in [BBF09]

$$\begin{aligned}\boldsymbol{\Sigma}_h &:= \mathcal{RT}_1(\mathcal{T}_h)^d\,,\\ \mathbf{U}_h &:= \mathcal{DP}_1(\mathcal{T}_h)^d\,,\\ \boldsymbol{\Theta}_h &:= \mathcal{P}_1(\mathcal{T}_h)^{d\times d} \cap \boldsymbol{\Theta}\end{aligned} \tag{2.22}$$

and point out that for $\mathbf{g}_N \in \mathcal{DP}(\mathcal{S}_{h,N})$ we have $\boldsymbol{\Sigma}_{\mathbf{g}_N,h} = \boldsymbol{\Sigma}_h \cap \boldsymbol{\Sigma}_{\mathbf{g}_N} \neq \emptyset$. Since we do not only have $\boldsymbol{\Sigma}_h \subset \boldsymbol{\Sigma}$ but also

$$\{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_{\mathbf{0},h} : (\operatorname{div}\boldsymbol{\tau}_h, \mathbf{v}_h) = 0\ \forall\, \mathbf{v}_h \in \mathbf{U}_h\} \subset \{\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}} : \operatorname{div}\boldsymbol{\tau} = \mathbf{0}\}\,, \tag{2.23}$$

together with (2.21) and (2.18) the standard theory of mixed finite elment methods (see [BBF13]) directly provides the following convergence result for our problem:

Theorem 2.3. *Let $(\boldsymbol{\sigma}, \mathbf{u}, \boldsymbol{\theta})$ be the sufficiently regular solution of (2.10). Let $(\boldsymbol{\Sigma}_h, \mathbf{U}_h, \boldsymbol{\Theta}_h)$ be defined as in (2.22), then (2.19) has a unique solution $(\boldsymbol{\sigma}_h, \mathbf{u}_h, \boldsymbol{\theta}_h)$ that satisfies*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\boldsymbol{\Sigma}} + \|\mathbf{u} - \mathbf{u}_h\|_{\mathbf{U}} + \|\boldsymbol{\theta} - \boldsymbol{\theta}_h\|_{\boldsymbol{\Theta}} \leq Ch^2 \tag{2.24}$$

with C independent of h.

Proof. See [BBF09]. □

2.4 Reconstruction-based a-posteriori error estimation

There exist several approaches to a-posteriori error estimation for linear elasticity both for the primal and the dual formulation (see for example [CD98, LV04, Kim11, BMS19]). The Least Squares formulation [CS03, CS04]) even comes with an already built in error estimator. Our goal is to derive an estimator that can be generalized in a rather straight forward manner to frictional contact problems. Consequently the discussion in this section is mainly to be understood as part of that derivation. The idea behind our approach can be traced back to [PS47] and [HHNL88] and exploits the strong duality that holds also for frictional contact problems (as we shall see in chapter 4):

For the exact solution

$$\mathcal{J}_{pr}(\mathbf{u}) + \mathcal{J}_{du}(\boldsymbol{\sigma}) = 0 \tag{2.25}$$

holds. Indeed we have

$$\begin{aligned}
\mathcal{J}_{pr}(\mathbf{u}) + \mathcal{J}_{du}(\boldsymbol{\sigma}) &= \frac{1}{2}(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{u}))_\Omega - (\mathbf{f}, \mathbf{u})_\Omega - (\mathbf{g}_N, \mathbf{u})_{\Gamma_N} \\
&\quad + \frac{1}{2}(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\sigma})_\Omega - \langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma + (\mathbf{g}_N, \mathbf{g}_D)_{\Gamma_N} \\
&= (\boldsymbol{\sigma}, \boldsymbol{\varepsilon}(\mathbf{u}))_\Omega - (\mathbf{f}, \mathbf{u})_\Omega - (\mathbf{g}_N, \mathbf{u}-\mathbf{g}_D)_{\Gamma_N} - \langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{g}_D\rangle_\Gamma \\
&= \langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{u}-\mathbf{g}_D\rangle_\Gamma - (\operatorname{div}\boldsymbol{\sigma} + \mathbf{f}, \mathbf{u})_\Omega - (\mathbf{g}_N, \mathbf{u}-\mathbf{g}_D)_{\Gamma_N} \\
&= 0\,.
\end{aligned} \tag{2.26}$$

This suggests that the violation of (2.25), also called "duality gap", can be used to measure the quality of the approximations $\boldsymbol{\sigma}_h$ and $\mathbf{u}_h$. An essential ingredient for the successful calculation in (2.26) is the validity of the material law $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{u})$, which represents the link between the dual and the primal variable and which will be the starting point of the derivation of our error estimator.

2.4.1 Reconstruction

Since the displacement approximation $\mathbf{u}_h$ from our saddle point formulation (2.19) is only in $\mathbf{L}^2(\Omega)$, a series of reconstruction steps is necessary to obtain a displacement approximation $\mathbf{u}_h^R \in \mathbf{V}_{\mathbf{g}_D}$ which we can use for a-posteriori error estimation. We basically adopt the approach in [Voh10] which itself relies on the procedure proposed in [Ste91]. Starting point is the observation that (2.15) implies $\boldsymbol{\nabla}\mathbf{u} = \mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}$ for the exact solution. Thus $\mathbf{G}_h := \mathcal{A}\boldsymbol{\sigma}_h + \boldsymbol{\theta}_h$ is a piecewise polynomial approximation of the gradient of displacements. The displacement reconstruction is done in the following steps:

(i) For each $T \in \mathcal{T}_h$, determine $\mathbf{u}_h^\circ|_T \in \mathcal{P}_k(T)^d$, such that

$$\begin{aligned}
(\boldsymbol{\nabla}\mathbf{u}_h^\circ, \boldsymbol{\nabla}\mathbf{v}_h)_{0,T} &= (\mathbf{G}_h, \boldsymbol{\nabla}\mathbf{v}_h)_{0,T} \quad \forall\, \mathbf{v}_h \in \mathcal{P}_k(T)^d\,, \\
(\mathbf{u}_h^\circ, \mathbf{e})_{0,T} &= (\mathbf{u}_h, \mathbf{e})_{0,T} \text{ for all } \mathbf{e} \in \mathcal{P}_0(T)^d\,.
\end{aligned} \tag{2.27}$$

(ii) Construct a conforming reconstruction $\mathbf{u}_h^R \in \mathcal{P}_k(\mathcal{T}_h)^d \subset \mathbf{H}^1(\Omega)$ by averaging:

$$\mathbf{u}_h^R(\mathbf{z}) = \frac{1}{\#\{T : \mathbf{z} \in T\}} \sum_{T:\mathbf{z}\in T} \mathbf{u}_h^\circ|_T(\mathbf{z}) \,, \tag{2.28}$$

where # denotes the cardinality of a set and $\mathbf{z}$ denote the nodes of the Lagrangian grid on the triangulation $\mathcal{T}_h$ (e.g. for $k = 2$ the vertices and edge midpoints).

(iii) Enforce the boundary conditions by setting

$$\mathbf{u}_h^R(\mathbf{z}) = \mathbf{g}_D(\mathbf{z}) \text{ for all } \mathbf{z} \in \Gamma_D \,. \tag{2.29}$$

Thus for $\mathbf{g}_D \in \mathcal{P}_k(\mathcal{S}_{h,D})$ we have $\mathbf{u}_h^R \in \mathbf{V}_{\mathbf{g}_D}$.

2.4.2 Reliability

Theorem 2.4. *Let $\boldsymbol{\sigma}$ solve (2.10) and $\boldsymbol{\sigma}_h$ solve (2.19). Let $\mathbf{f} \in \mathcal{DP}_1(\mathcal{T}_h)$ and $\mathbf{u}_h^R \in \mathbf{V}_{\mathbf{g}_D}$. Then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathrm{div})} \le C\,(\eta_1 + \eta_2 + \eta_3) \tag{2.30}$$

holds with

$$\begin{aligned}
\eta_1 &:= \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|_0 \,, \\
\eta_2 &:= \|\mathbf{as}\,\boldsymbol{\sigma}_h\|_0 \,, \\
\eta_3 &:= \|\frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma}_h - \mathrm{div}\,\mathbf{u}_h^R\|_0
\end{aligned}$$

and a constant C which is independent of λ and h.

Proof. First we observe that since $\mathbf{f} \in \mathcal{DP}_1(\mathcal{T}_h)$ we have $\mathrm{div}\,\boldsymbol{\sigma}_h = \mathbf{f}$ and consequently $\mathrm{div}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = 0$ and

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathrm{div})} = \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_0 \,. \tag{2.31}$$

Starting with η_1^2 and inserting $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{u})$ leads to

$$\begin{aligned}
\eta_1^2 &= \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|^2 = \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|^2 \\
&= \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 + \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|^2 - 2(\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)) \\
&= \|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 + \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|^2 - \frac{1}{\mu}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)) \\
&\quad + \frac{\lambda}{\mu(d\lambda + 2\mu)}(\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \mathrm{div}(\mathbf{u} - \mathbf{u}_h^R)) \,,
\end{aligned} \tag{2.32}$$

where we have used the specific form of $\mathcal{A}$ from (2.4). For the first mixed term in (2.32), integration by parts leads to

$$\begin{aligned}
(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)) &= (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\nabla}(\mathbf{u} - \mathbf{u}_h^R)) - (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{as}\,\boldsymbol{\nabla}(\mathbf{u} - \mathbf{u}_h^R)) \\
&= \langle(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\cdot\mathbf{n}, \mathbf{u} - \mathbf{u}_h^R\rangle_\Gamma + (\mathbf{as}\,\boldsymbol{\sigma}_h, \boldsymbol{\nabla}(\mathbf{u} - \mathbf{u}_h^R)) \\
&\le C_K\|\mathbf{as}\,\boldsymbol{\sigma}_h\|\|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\| \,,
\end{aligned} \tag{2.33}$$

where we have used the Cauchy-Schwarz and Korn's inequalities (1.54) as well as the fact that since $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{\Sigma_0}$ and $\mathbf{u} - \mathbf{u}_h^R \in \mathbf{V}$ the boundary term vanishes. The second mixed term in (2.32) may be treated as

$$\begin{aligned}(\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \mathrm{div}(\mathbf{u} - \mathbf{u}_h^R)) &= (\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma} - \mathrm{div}\,\mathbf{u}_h^R)) \\ &= \frac{1}{d\lambda + 2\mu}\|\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 \\ &\quad + (\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma}_h - \mathrm{div}\,\mathbf{u}_h^R) \\ &\geq (\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma}_h - \mathrm{div}\,\mathbf{u}_h^R)\,.\end{aligned}$$

Thus applying again the Cauchy-Schwarz inequality and (1.56) we obtain

$$-(\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \mathrm{div}(\mathbf{u} - \mathbf{u}_h^R)) \leq C_D\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|\|\frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma}_h - \mathrm{div}\,\mathbf{u}_h^R\|. \tag{2.34}$$

Combining (2.32)-(2.34) and applying Young's inequality (1.57) twice leads to

$$\begin{aligned}\mu\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 + \mu\|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|^2 \leq \mu\|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|^2 \\ + \frac{1}{2\delta_1}\|\mathbf{as}\,\boldsymbol{\sigma}_h\|^2 + \frac{\delta_1}{2}C_K^2\|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|^2 \\ + \frac{\delta_2}{2}\left(\frac{\lambda C_D}{d\lambda + 2\mu}\right)^2\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 \\ + \frac{1}{2\delta_2}\|\mathrm{div}\,\mathbf{u}_h^R - \frac{1}{d\lambda + 2\mu}\mathrm{tr}\,\boldsymbol{\sigma}_h)\|^2\end{aligned}$$

with positive numbers δ_1, δ_2 to be chosen appropriately. The observation that (1.56) implies

$$\begin{aligned}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 &= \|2\mu\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + \frac{\lambda}{d\lambda + 2\mu}\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\mathbf{I}\|^2 \\ &\leq 8\mu^2\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 + 2d\left(\frac{\lambda}{d\lambda + 2\mu}\right)^2\|\mathrm{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2 \\ &\leq \left(8\mu^2 + 2d\left(\frac{\lambda}{d\lambda + 2\mu}\right)^2 C_D\right)\|\mathcal{A}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2\,,\end{aligned}$$

and choosing δ_1, δ_2 sufficiently small finishes the proof. □

Remark. The requirement that $\boldsymbol{\sigma}_h$ solves (2.19) is not absolutely necessary. In fact it can be replaced by any $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h^*$ in the discrete admissible set.

Remark. η_2 can be omitted because of

$$\|\mathbf{as}\,\boldsymbol{\sigma}_h\| = 2\mu\|\mathbf{as}\,\mathcal{A}\boldsymbol{\sigma}_h\| = 2\mu\|\mathbf{as}(\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R))\| \leq 2\mu\|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|\,. \tag{2.35}$$

However, it might be useful in some settings to monitor the violation of the symmetry condition explicitly.

Remark. For compressible materials ($\lambda < \infty$), η_1 and η_3 can be contracted into the energy norm of the residual and we have the concise estimate

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 \leq \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|_{\mathcal{C}}^2 + C_K^2 \|\mathbf{as}\, \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 \tag{2.36}$$

with

$$\begin{aligned}
\|\boldsymbol{\tau}\|_{\mathcal{C}}^2 &:= (\mathcal{C}\boldsymbol{\tau}, \boldsymbol{\tau}) = 2\mu \|\boldsymbol{\tau}\|_0^2 + \lambda \|\operatorname{tr} \boldsymbol{\tau}\|_0^2 \\
\|\boldsymbol{\tau}\|_{\mathcal{A}}^2 &:= (\mathcal{A}\boldsymbol{\tau}, \boldsymbol{\tau}) = \frac{1}{2\mu} \|\boldsymbol{\tau}\|_0^2 - \frac{\lambda}{2\mu(d\lambda + 2\mu)} \|\operatorname{tr} \boldsymbol{\tau}\|_0^2 .
\end{aligned}$$

The estimate in the energy norm (2.36) then implies an estimate similar to (2.30) but with a constant that is in general dependent on λ:

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\operatorname{div})} \leq C(\lambda) \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|_{\mathcal{C}} . \tag{2.37}$$

Proof.

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 &= \|\boldsymbol{\varepsilon}(\mathbf{u}_h^R) - \mathcal{A}\boldsymbol{\sigma}_h\|_{\mathcal{C}}^2 - \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|_{\mathcal{C}}^2 + 2(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)) \\
&\leq \|\boldsymbol{\varepsilon}(\mathbf{u}_h^R) - \mathcal{A}\boldsymbol{\sigma}_h\|_{\mathcal{C}}^2 - \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|_{\mathcal{C}}^2 \\
&\quad + 2C_K \|\mathbf{as}\, \boldsymbol{\sigma}_h\| \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R))\| \\
&\leq \|\boldsymbol{\varepsilon}(\mathbf{u}_h^R) - \mathcal{A}\boldsymbol{\sigma}_h\|_{\mathcal{C}}^2 - \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R)\|_{\mathcal{C}}^2 \\
&\quad + C_K^2 \|\mathbf{as}\, \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 + \|\boldsymbol{\varepsilon}(\mathbf{u} - \mathbf{u}_h^R))\|_{\mathcal{C}}^2 \\
&\leq \|\boldsymbol{\varepsilon}(\mathbf{u}_h^R) - \mathcal{A}\boldsymbol{\sigma}_h\|_{\mathcal{C}}^2 + C_K^2 \|\mathbf{as}\, \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 .
\end{aligned}$$

(2.37) results from (2.36) with (2.31) and the following easily verifiable inequalities

$$\begin{aligned}
\|\boldsymbol{\tau}\|_{\mathcal{A}}^2 &\geq \frac{1}{d\lambda + 2\mu} \|\boldsymbol{\tau}\|_0^2 , \\
\|\mathbf{as}\, \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 &\leq \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|_{\mathcal{C}}^2 .
\end{aligned}$$

□

Remark. The global Korn's constant C_K can be replaced by a local one, by exploiting the weak symmetry of $\boldsymbol{\sigma}_h$ (see [BKMS18]).

Remark. In the case of contact problems the boundary term in (2.33) no longer vanishes. Its appropriate treatment will be discussed thoroughly in the respective sections of chapters 3 and 4.

Since all terms of the estimator are L^2-norms the decomposition with respect to the triangulation is straightforward:

$$\begin{aligned}
\eta_{1,T} &:= \|\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\varepsilon}(\mathbf{u}_h^R)\|_{L^2(T)} , \\
\eta_{2,T} &:= \|\mathbf{as}\, \boldsymbol{\sigma}_h\|_{L^2(T)} , \\
\eta_{3,T} &:= \|\frac{1}{d\lambda + 2\mu} \operatorname{tr} \boldsymbol{\sigma}_h - \operatorname{div} \mathbf{u}_h^R\|_{L^2(T)} , \\
\eta_i^2 &= \sum_{T \in \mathcal{T}_h} \eta_{i,T}^2 \quad \text{for } i = 1, 2, 3 .
\end{aligned}$$

The efficiency (as part of an adaptive refinement scheme) of the a-posteriori error estimator discussed in this section will be tested by a number of numerical results in the following section.

2.5 Numerical Experiments

In this section we report some results achieved by our implementation of the discussed finite element method combined with the reconstruction based error estimator and the resulting adaptive mesh refinement strategy. A Dörfler marking strategy is applied, which consists of finding the smallest set of triangles $\tilde{\mathcal{T}}_h \subset \mathcal{T}_h$ such that

$$\sum_{T \in \tilde{\mathcal{T}}_h} \eta_T^2 \geq \theta^2 \sum_{T \in \mathcal{T}_h} \eta_T^2$$

holds for a chosen paramter θ. All triangles in this set are then refined as well as those adjacent triangles necessary to avoid hanging nodes. All computations (in this and the following chapters) were carried out in MATLAB.

Cook's Membrane

We will consider the well known Cook's Membrane problem as a benchmark problem to test the performance of the presented method. The reference configuration is given by $\Omega \in \mathbb{R}^2$, which is the quadrilateral defined as the convex hull of its corners $\{(0,0),(0.48,0.44),(0.48,0.6),(0,0.44)\}$. Displacement $\mathbf{g}_D$ on the Dirichlet boundary $\Gamma_D = \{0\} \times (0, 0.44)$ as well as volume forces $\mathbf{f}$ are set to zero. On the Neumann boundary $\Gamma_N = \mathrm{int}(\Gamma \setminus \overline{\Gamma_D})$ the following surface traction forces are applied:

$$\mathbf{g}_N(\mathbf{x}) := \begin{cases} \begin{pmatrix} 0 \\ 0.05 \end{pmatrix} & \mathbf{x} \in \{0.48\} \times (0.44, 0.6) \\ \mathbf{0} & \text{else} \end{cases}$$

While the shear modulus μ is scaled to 1 in all cases, we will test both the compressible case with $\lambda = 1$ and the fully incompressible case ($\lambda = \infty$).

Table 2.1 shows the results of the compressible case on a sequence of adaptively refined triangulations obtained with Dörfler parameter $\theta = 0.8$. Here and in the following η is defined by $\eta^2 := \sum_i \eta_i^2$ and $|\mathcal{T}_h|$ denotes the number of triangles in the triangulation $\mathcal{T}_h$. Table 2.2 reports the incompressible case for the same Dörfler parameter.

Figure 2.1 depicts the initial and deformed configurations of the compressible case after 7 steps of adaptive refinement. The deformed configuration is obtained using the reconstructed displacement $\mathbf{u}_h^R \in \mathcal{P}_2(\mathcal{T}_h)$. As expected, refinement concentrates at the corners, particularly in the upper left corner where the solution is known to feature a singularity. This can be seen in Figure 2.2, where we plotted the von-Mises stress given by the following formula.

$$\sigma_{VM} = \sqrt{\sigma_{11}^2 + \sigma_{22}^2 - \sigma_{11}\sigma_{22} + \frac{3}{2}(\sigma_{12}^2 + \sigma_{21}^2)}$$

Visualizing the distribution of the components of the stress tensor in figure 2.3, we observe that the weak symmetry condition is strong enough to obtain qualitatively indistinguishable off-diagonal entries.

The behaviour of the different parts of the error estimator for the compressible case (the incompressible case is comparable) is depicted in Figure 2.4. Since $\mu = 1$ the difference of the three contributions is minor. For $\mu \neq 1$, we observed $\eta_2 \approx \mu\eta_1$ which is in accordance with (2.35).

A comparison of the reduction of η for uniform and adaptive refinement is given in Figure 2.5. While uniform refinement yields only a suboptimal rate due to the lack of regularity of the solution, the optimal convergence behaviour $\eta \sim N_h^{-1}$ (for N_h being the total number of degrees of freedom) is recovered by our adaptive refinement procedure, illustrating the efficiency of the proposed error estimator.

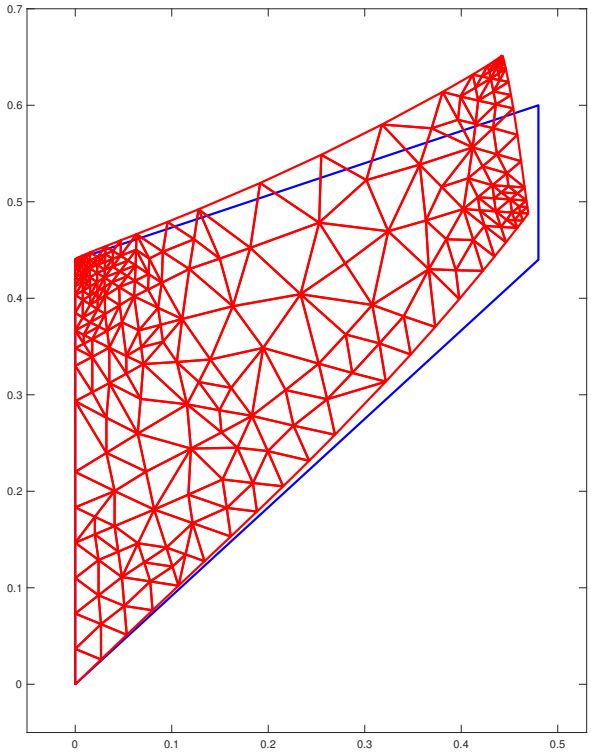

Figure 2.1: Deformed mesh after 7 refinements for $\lambda = \infty$.

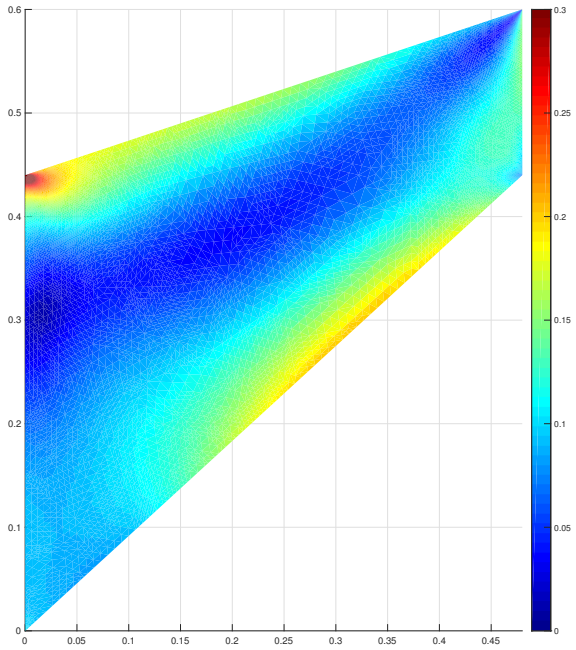

Figure 2.2: Von-Mises stress distribution for $\lambda = \infty$.

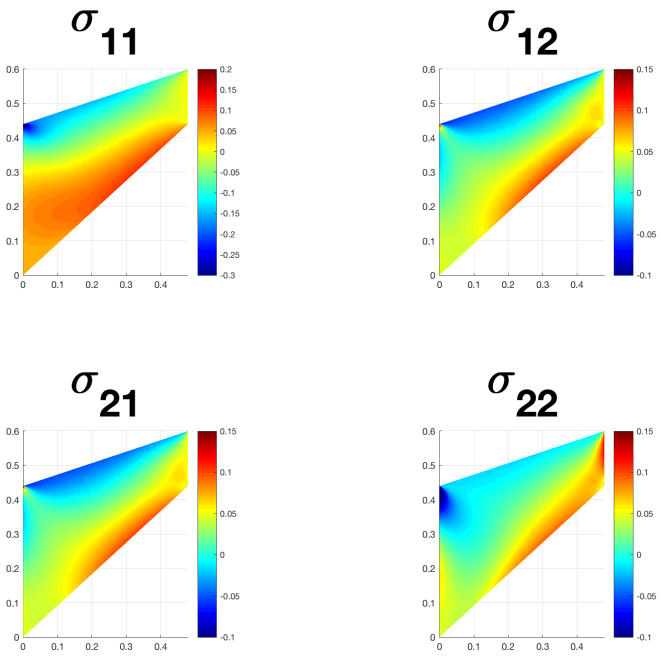

Figure 2.3: Stress distribution for $\lambda = \infty$.

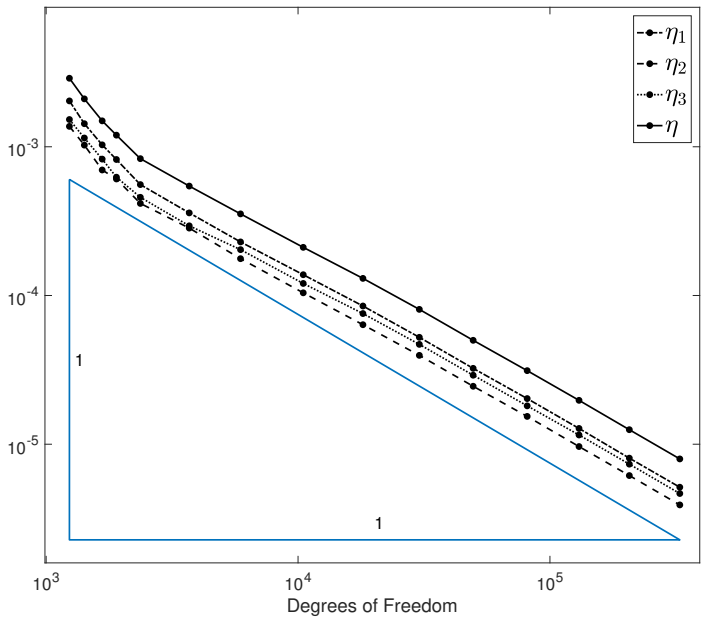

Figure 2.4: Convergence of error estimator terms for $\lambda = 1$.

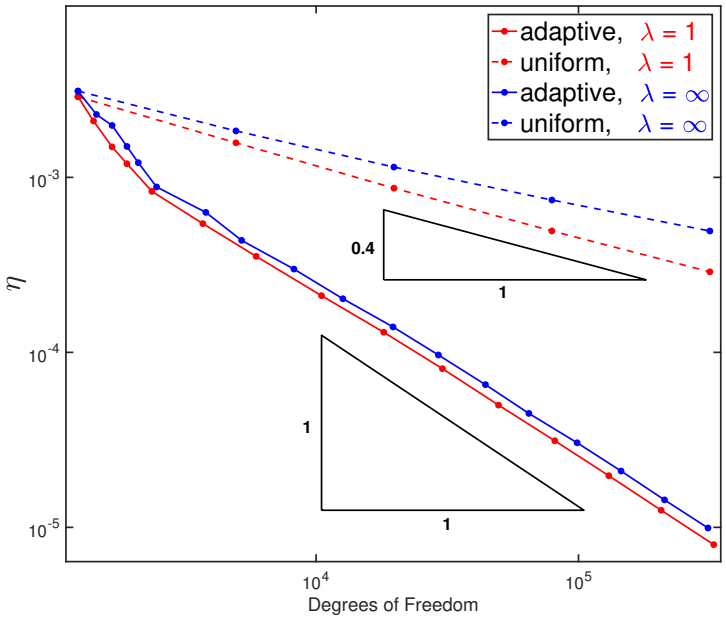

Figure 2.5: Adaptive vs. uniform refinement.

l	$\lvert\mathcal{T}_h\rvert$	$\dim \boldsymbol{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \boldsymbol{\Theta}_h$	η_1	η_2	η_3	η	eoc
0	75	724	450	51	0.00203412	0.00152494	0.00137121	0.00288847	
1	86	828	516	58	0.00142934	0.00114524	0.00102582	0.00209926	2.34
2	101	976	606	67	0.00103014	0.000826141	0.000698147	0.00149369	2.09
3	115	1116	690	75	0.000821021	0.000624327	0.000607374	0.00119698	1.7
4	143	1388	858	92	0.000556768	0.000457203	0.000415865	0.000831847	1.66
5	223	2180	1338	137	0.000359222	0.000294103	0.000283931	0.000544201	0.95
6	357	3504	2142	213	0.000229195	0.000203376	0.00017682	0.000353776	0.92
7	634	6248	3804	363	0.000137942	0.000120504	0.000104241	0.000210749	0.9
8	1095	10824	6570	610	8.49861e-05	7.56157e-05	6.36188e-05	0.000130337	0.88
9	1834	18176	11004	1000	5.23359e-05	4.6921e-05	3.95812e-05	8.06678e-05	0.93
10	3000	29792	18000	1604	3.23736e-05	2.91059e-05	2.45308e-05	4.99697e-05	0.97
11	4918	48900	29508	2599	2.02745e-05	1.81136e-05	1.53834e-05	3.12379e-05	0.95
12	7900	78648	47400	4126	1.27668e-05	1.15392e-05	9.64745e-06	1.97286e-05	0.97
13	12515	124680	75090	6485	8.06093e-06	7.34518e-06	6.15588e-06	1.2523e-05	0.99
14	19852	197932	119112	10216	5.13923e-06	4.66306e-06	3.90825e-06	7.96431e-06	0.98

Table 2.1: Results for compressible case ($\lambda = 1$)

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	η_1	η_2	η_3	η	eoc
0	75	724	450	51	0.0022052	0.00162469	0.00147519	0.00311107	
1	88	852	528	59	0.00160578	0.00134835	0.000913143	0.00228701	1.9
2	101	976	606	67	0.00134086	0.00120489	0.00081922	0.0019801	1.05
3	115	1116	690	75	0.00101185	0.000871366	0.000687755	0.00150204	2.12
4	127	1236	762	82	0.000808653	0.000664363	0.000613966	0.00121336	2.14
5	149	1448	894	96	0.000574644	0.000500068	0.000443358	0.00088139	1.99
6	229	2240	1374	141	0.000404427	0.000358083	0.000325801	0.000630818	0.77
7	314	3080	1884	189	0.00028005	0.000250105	0.000222484	0.00043644	1.17
8	497	4904	2982	289	0.000191848	0.000171513	0.000152569	0.000299165	0.82
9	764	7552	4584	434	0.000129433	0.000116985	0.0001027	0.000202449	0.91
10	1188	11744	7128	663	8.90646e-05	8.03545e-05	7.15961e-05	0.000139697	0.84
11	1770	17540	10620	968	6.14899e-05	5.55172e-05	4.9654e-05	9.65851e-05	0.93
12	2674	26548	16044	1443	4.15242e-05	3.77432e-05	3.35877e-05	6.53983e-05	0.94
13	3917	38900	23502	2089	2.84358e-05	2.57828e-05	2.30753e-05	4.47863e-05	0.99
14	5985	59548	35910	3149	1.93548e-05	1.75022e-05	1.56401e-05	3.04229e-05	0.91
15	8799	87632	52794	4593	1.32887e-05	1.20834e-05	1.08488e-05	2.09832e-05	0.96
16	12883	128372	77298	6683	9.07727e-06	8.3034e-06	7.39198e-06	1.43522e-05	1
17	18873	188180	113238	9723	6.28033e-06	5.74376e-06	5.08911e-06	9.91627e-06	0.97

Table 2.2: Results for incompressible case ($\lambda = \infty$)

Chapter 3

A dual formulation of the Signorini Problem

In this chapter we consider the (frictionless) Signorini problem which can be seen as a generalization of the problem of linear elasticity. Since it is concerned with only one elastic object that is in contact with a rigid obstacle, it is sometimes referred to as one-sided contact problem and the effect of the contact is represented by an unilateral constraint. The general setting is illustrated in Figure 3.1. While in the literature on the Signorini problem the focus lies mostly on the displacement approach, we will concentrate on the stress-based formulation in order to be able to generalize the finite elment method that was discussed in the last chapter.

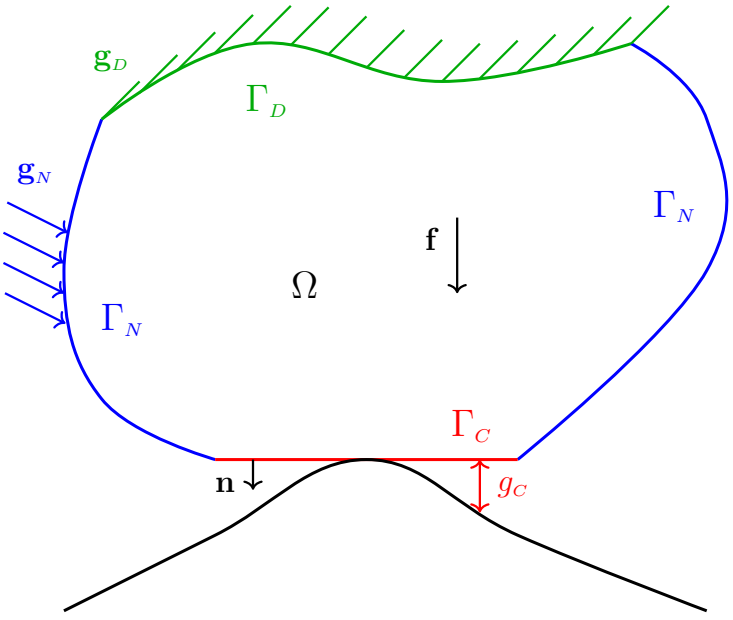

Figure 3.1: General setting of the Signorini problem.

3.1 An unilateral Contact Problem

As mentioned in chapter 2 the classical formulation of the Signorini problem is identical to the problem of linear elasticity with the exception that the boundary conditions are changed to reflect the effect of the obstacle. For this purpose the boundary is divided into three disjoint parts $\partial\Omega = \Gamma = \overline{\Gamma_D \uplus \Gamma_N \uplus \Gamma_C}$. Γ_C is the part of the boundary where contact may possibly occur, and it is assumed that it is sufficiently smooth and surrounded by Γ_N (see section 1.1.3 for details). The distance in normal direction between Γ_C and the obstacle is represented by the gap-function g_C. The contact conditions can then be expressed by the following equations:

$$g_C - \mathbf{u}_n \geq 0 \qquad \text{on } \Gamma_C\,, \tag{3.1a}$$
$$(\boldsymbol{\sigma}\cdot\mathbf{n})_n \leq 0 \qquad \text{on } \Gamma_C\,, \tag{3.1b}$$
$$(\boldsymbol{\sigma}\cdot\mathbf{n})_n(g_C - \mathbf{u}_n) = 0 \qquad \text{on } \Gamma_C\,, \tag{3.1c}$$
$$(\boldsymbol{\sigma}\cdot\mathbf{n})_t = 0 \qquad \text{on } \Gamma_C\,. \tag{3.1d}$$

The first inequality (3.1a) is called the *non-penetration condition* since it forbids the normal displacement to exceed the distance to the obstacle and thus penetrate it. Wherever $\mathbf{u}_n = g_C$ holds on Γ_C the elastic solid is in contact with the obstacle. The second inequality (3.1b) is an expression of the assumption that the obstacle is not adhesive and thus any contact with it will only result in compressive forces acting on the elastic solid. It is sometimes called *sign condition*. (3.1c) is called the *complementarity condition* and summarizes the fact that there can only be forces acting on the elastic solid if it is in contact with the obstacle. Finally (3.1d) is called the *frictionless condition* and will be replaced in chapter 4 by the Coulomb friction law. Again we summarize the *classical frictionless problem* before turning to the corresponding variational formulations:

Find sufficiently smooth $\mathbf{u}$ and $\boldsymbol{\sigma}$ such that

$$\begin{aligned}
\operatorname{div}\boldsymbol{\sigma} &= -\mathbf{f} && \text{in } \Omega\,,\\
\mathbf{as}\,\boldsymbol{\sigma} &= \mathbf{0} && \text{in } \Omega\,,\\
\mathcal{A}\boldsymbol{\sigma} &= \boldsymbol{\varepsilon}(\mathbf{u}) && \text{in } \Omega\,,\\
\boldsymbol{\sigma}\cdot\mathbf{n} &= \mathbf{g}_N && \text{on } \Gamma_N\,,\\
\mathbf{u} &= \mathbf{g}_D && \text{on } \Gamma_D\,,\\
g_C - \mathbf{u}_n &\geq 0 && \text{on } \Gamma_C\,,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n &\leq 0 && \text{on } \Gamma_C\,,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n(g_C - \mathbf{u}_n) &= 0 && \text{on } \Gamma_C\,,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_t &= 0 && \text{on } \Gamma_C\,.
\end{aligned} \tag{3.2}$$

3.2 Variational formulations

The variational formulations introduced in section 2.2 can be extended in a rather straigthforward manner by incorporating the respective contact conditions into the ansatz spaces of the minimization problems.

Stress-based formulation

The *dual frictionless problem* consists in finding $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{K}}_0^*$ such that

$$\mathcal{J}_{du}^C(\boldsymbol{\sigma}) = \min_{\boldsymbol{\tau} \in \boldsymbol{\mathcal{K}}_0^*} \mathcal{J}_{du}^C(\boldsymbol{\tau}) := \frac{1}{2}(\mathcal{A}\boldsymbol{\tau}, \boldsymbol{\tau})_\Omega - \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma - \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} . \tag{3.3}$$

Here the conditions (3.1b) and (3.1d) are explicitly included in the admissible set $\boldsymbol{\mathcal{K}}_0^*$ while the non-penetration and complementarity conditions are hidden in the structure of the functional. Due to $\boldsymbol{\mathcal{K}}_0^*$ being only a closed convex subset of $\boldsymbol{\Sigma}$ (not a subspace), (3.3) is equivalent to a variational inequality (instead of an equality):

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau} - \boldsymbol{\sigma})_\Omega \geq \langle (\boldsymbol{\tau} - \boldsymbol{\sigma}) \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle ((\boldsymbol{\tau} - \boldsymbol{\sigma}) \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\mathcal{K}}_0^* . \tag{3.4}$$

Recalling the definition of $\boldsymbol{\mathcal{K}}_0^*$ in (1.25) we can again enforce the constraints with Lagrange multipliers and write down a KKT-System for (3.3):

Find $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_N + \hat{\boldsymbol{\sigma}}) \in (\boldsymbol{\sigma}_N + \boldsymbol{\Sigma}_{\mathbf{0}}^C)$, $\mathbf{u} \in \mathbf{U}$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $p \in P$ such that

$$\begin{aligned}
(\mathcal{A}\hat{\boldsymbol{\sigma}}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div} \boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\, \boldsymbol{\tau}) + & \\
\langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} &= \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} - (\mathcal{A}\boldsymbol{\sigma}_N, \boldsymbol{\tau}) , \\
(\operatorname{div} \hat{\boldsymbol{\sigma}}, \mathbf{v}) &= -(\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_N, \mathbf{v}) , \\
(\mathbf{as}\, \hat{\boldsymbol{\sigma}}, \boldsymbol{\gamma}) &= -(\mathbf{as}\, \boldsymbol{\sigma}_N, \boldsymbol{\gamma}) , \\
\langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, w \rangle_{\Gamma_C} &\leq 0 , \\
\langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} &= 0
\end{aligned} \tag{3.5}$$

holds for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}^C$, $\mathbf{v} \in \mathbf{U}$, $\boldsymbol{\gamma} \in \boldsymbol{\Theta}$ and $w \in P$.

Remark. Do not confuse $p \in P$ in (3.5) with the pressure from the incompressible displacement-pressure formulation in (2.9).

Displacement formulation

Let $\mathbf{u}_D \in \mathbf{V}_{\mathbf{g}_D}$. Then the *primal frictionless problem* consists in finding $\mathbf{u} \in (\mathbf{u}_D + \mathbf{K}_{g_C})$ such that

$$\mathcal{J}_{pr}(\mathbf{u}) = \min_{\mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C})} \mathcal{J}_{pr}(\mathbf{v}) := \frac{1}{2}(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{v}), \boldsymbol{\varepsilon}(\mathbf{v}))_\Omega - (\mathbf{f}, \mathbf{v})_\Omega - (\mathbf{g}_N, \mathbf{v})_{\Gamma_N} , \tag{3.6}$$

which is also equivalent to a variational inequality:

$$(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega \geq (\mathbf{f}, \mathbf{v} - \mathbf{u})_\Omega + (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} \quad \forall\, \mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C}) . \tag{3.7}$$

In contrast to the stress-based formulation, the displacement formulation explicitly requires the non-penetration (3.1a) condition to hold, while the other contact conditions are enforced implicitly.

3.2.1 Duality

Theorem 3.1.

(i) Let $\mathbf{u}$ *be the solution of (3.7), then* $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u})$, $\mathbf{u}$, $\boldsymbol{\theta} := \mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}$ *and* $p := g_C - u_n$ *solve (3.5).*

(ii) Let conversely $\boldsymbol{\sigma}$, $\mathbf{u}$, $\boldsymbol{\theta}$ *and* p *be the solution of (3.5), then* $\mathbf{u}$ *is in* $H^1(\Omega)$ *and solves (3.7).*

Proof. We restrict the proof to the case $\lambda < \infty$ since the treatment of the incompressible case is analogous to the proof of the remark after (2.1).

(i) The first part of the proof is also discussed in [KO88, Chapter 6]. We repeat it here in our notation. Let $\mathbf{u}$ be the solution of (3.7). Then $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}) \in L^2(\Omega)^{d\times d}$ and $\mathbf{as}\,\boldsymbol{\sigma} = \mathbf{0}$. Testing with $\mathbf{v} = \mathbf{u} \pm \boldsymbol{\phi}$ with $\boldsymbol{\phi} \in C_c^\infty(\Omega)^d$ in (3.7) we obtain $\operatorname{div}\boldsymbol{\sigma} = -\mathbf{f}$ implying $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^{\mathbf{f}}$ and

$$\langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{v}-\mathbf{u}\rangle_{\Gamma} \geq (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N} \quad \forall\, \mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C})\,,$$

which is equivalent to

$$ {}_{00}\langle \boldsymbol{\sigma}\cdot\mathbf{n}, \hat{\mathbf{v}}-\hat{\mathbf{u}}\rangle_{\Gamma_R} \geq (\mathbf{g}_N, \hat{\mathbf{v}}-\hat{\mathbf{u}})_{\Gamma_N} \quad \forall\, \hat{\mathbf{v}} \in \mathbf{H}_{00}^{1/2}(\Gamma_R) \text{ s.t. } g_C - \hat{v}_n \in P \tag{3.8}$$

with $\hat{\mathbf{u}} := \mathbf{u} - \mathbf{u}_D$. Testing in (3.8) with $\hat{\mathbf{v}} = \hat{\mathbf{u}} \pm \boldsymbol{\phi}$ with $\boldsymbol{\phi} \in \mathbf{H}_{00}^{1/2}(\Gamma_R)$ such that $\operatorname{supp}(\boldsymbol{\phi}) \subset \Gamma_N$ yields $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^F_{\mathbf{g}_N}$, allowing us to write (3.8) as

$$\langle \hat{\boldsymbol{\sigma}}\cdot\mathbf{n}, \hat{\mathbf{v}}-\hat{\mathbf{u}}\rangle_{\Gamma_C} \geq 0 \quad \forall\, \hat{\mathbf{v}} \in \mathbf{H}^{1/2}(\Gamma_C) \text{ s.t. } g_C - \hat{v}_n \in P \tag{3.9}$$

with $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_N$. Testing in (3.9) with $\hat{\mathbf{v}} = \hat{\mathbf{u}} \pm \boldsymbol{\phi}$ with $\boldsymbol{\phi} \in \mathbf{H}^{1/2}(\Gamma_C)$ such that $\phi_n = 0$ yields $\hat{\boldsymbol{\sigma}} \in \boldsymbol{\Sigma}^C_{\mathbf{0}}$ and

$$\langle (\hat{\boldsymbol{\sigma}}\cdot\mathbf{n})_n, w + \hat{u}_n - g_C\rangle_{\Gamma_C} \leq 0 \quad \forall\, w \in P\,. \tag{3.10}$$

Finally testing in (3.10) with $w = 0$ and $w = 2(g_C - \hat{u}_n)$ and setting $p := g_C - u_n = g_C - \hat{u}_n$, we obtain the last two relations in (3.5).

What remains to be shown is the first equation in (3.5). Setting $\boldsymbol{\theta} := \mathbf{as}\,\nabla\mathbf{u}$, inserting all variables into the equation and applying partial integration leads to

$$\begin{aligned}
(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\,\boldsymbol{\tau}) + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, p\rangle_{\Gamma_C} &= \\
(\mathcal{A}\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div}\boldsymbol{\tau}) + (\mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}, \mathbf{as}\,\boldsymbol{\tau}) + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C - u_n\rangle_{\Gamma_C} &= \\
(\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\tau}) - (\boldsymbol{\nabla}\mathbf{u}, \boldsymbol{\tau}) + (\mathbf{as}\,\boldsymbol{\nabla}\mathbf{u}, \boldsymbol{\tau}) + \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{u}\rangle_{\Gamma} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C - u_n\rangle_{\Gamma_C} &= \\
\langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{u}_D\rangle_{\Gamma} + \langle \boldsymbol{\tau}\cdot\mathbf{n}, \hat{\mathbf{u}}\rangle_{\Gamma_C} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C - u_n\rangle_{\Gamma_C} &= \\
\langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_{\Gamma} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, u_n\rangle_{\Gamma_C} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C - u_n\rangle_{\Gamma_C} &= \\
\langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_{\Gamma} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C\rangle_{\Gamma_C} &
\end{aligned}$$

for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}^C$, where we used $\hat{\mathbf{u}} \in \mathbf{V}$ and $\mathcal{A} = \mathcal{C}^{-1}$.

(ii) Let now $\boldsymbol{\sigma}$, $\mathbf{u}$, $\boldsymbol{\theta}$ and p solve (3.5). By testing only with $\boldsymbol{\tau} \in C_c^\infty(\Omega)^{d\times d}$ in the first equation in (3.5) we obtain

$$(\mathbf{u}, \operatorname{div} \boldsymbol{\tau}) = -(\mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}, \boldsymbol{\tau}) \quad \forall\, \boldsymbol{\tau} \in C_c^\infty(\Omega)^{d\times d} \tag{3.11}$$

implying $\mathbf{u} \in H^1(\Omega)$ with $\boldsymbol{\nabla}\mathbf{u} = \mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}$. Consequently partial integration yields

$$\langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{u}\rangle_{\Gamma} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, p\rangle_{\Gamma_C} = \langle \boldsymbol{\tau}\cdot\mathbf{n}, \mathbf{g}_D\rangle_{\Gamma} + \langle (\boldsymbol{\tau}\cdot\mathbf{n})_n, g_C\rangle_{\Gamma_C} \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}}^C\,. \tag{3.12}$$

Testing only with $\boldsymbol{\tau} \in C^\infty(\Omega)^{d\times d}$ such that $\boldsymbol{\tau}$ vanishes on Γ_R yields $\mathbf{u} = \mathbf{g}_D$ on Γ_D. Setting again $\hat{\mathbf{u}} := \mathbf{u} - \mathbf{u}_D \in \mathbf{V}$ we obtain $g_C - \hat{u}_n = p$ on Γ_R and hence $\hat{\mathbf{u}} \in \mathbf{K}_{g_C}$. Finally observing $(\mathbf{u}_D + \mathbf{K}_{g_C}) - \mathbf{u} \subset \mathbf{V} \subset \mathbf{U}$ and $\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}) = \boldsymbol{\sigma}$ (which is a consequence of $\mathbf{as}\,\boldsymbol{\sigma} = \mathbf{0}$), partial integration of the second equation in (3.5) yields

$$\begin{aligned}
(\mathbf{f}, \mathbf{v}-\mathbf{u})_\Omega &= -(\operatorname{div}\boldsymbol{\sigma}, \mathbf{v}-\mathbf{u})_\Omega = (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - \langle \boldsymbol{\sigma}\cdot\mathbf{n}, \mathbf{v}-\mathbf{u}\rangle_{\Gamma} \\
&= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - \langle (\boldsymbol{\sigma}_N + \hat{\boldsymbol{\sigma}})\cdot\mathbf{n}, \mathbf{v}-\mathbf{u}\rangle_{\Gamma_R} \\
&= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N} - \langle \hat{\boldsymbol{\sigma}}\cdot\mathbf{n}, \hat{\mathbf{v}}-\hat{\mathbf{u}}\rangle_{\Gamma_C} \\
&= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N} \\
&\quad - \langle (\hat{\boldsymbol{\sigma}}\cdot\mathbf{n})_n, \hat{v}_n - g_C + g_C - \hat{u}_n\rangle_{\Gamma_C} \\
&= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N} \\
&\quad + \langle (\hat{\boldsymbol{\sigma}}\cdot\mathbf{n})_n, g_C - \hat{v}_n\rangle_{\Gamma_C} - \langle (\hat{\boldsymbol{\sigma}}\cdot\mathbf{n})_n, p\rangle_{\Gamma_C} \\
&\le (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N}
\end{aligned}$$

for all $\mathbf{v} = \mathbf{u}_D + \hat{\mathbf{v}}$ with $\hat{\mathbf{v}} \in \mathbf{K}_{g_C}$, where in the final step we used the last two relations in (3.5). □

3.2.2 Existence and Uniqueness

$\mathcal{J}_{du}^C(\boldsymbol{\sigma})$ differs from $\mathcal{J}_{du}(\boldsymbol{\sigma})$ only by a linear term and the inf-sup condition (2.17) also holds for $\boldsymbol{\mathcal{K}}_0^*$ because the boundary conditions are only weakened. Since $\boldsymbol{\mathcal{K}}_0^*$ is also convex, closed and nonempty, Theorem 1.12 again guarantees the unique solvability of (3.3) and its equivalent formulation as a variational inequality (3.4). Since the same arguments hold for the primal frictionless problem (3.6), Theorem 3.1 yields the unique existence of the Lagrange multipliers and thus the sufficiency and necessity of (3.5). We again summarize these observations in the following theorem:

Theorem 3.2. *Both the dual frictionless problem (3.3) or equivalently (3.5) and the primal frictionless problem (3.6) or equivalently (3.7) admit a unique solution.*

3.3 Finite Element Discretization

The starting point for our discretization of (3.5) will of course be (2.19) with the spaces defined in (2.22). In order to be able to concentrate on the issues arising from contact

we will subsequently assume that the given volume and traction forces $\mathbf{f}$ and $\mathbf{g}_N$ are piecewise affine and compatible with the triangulation.

While the treatment of standard boundary conditions such as on Γ_N is rather straightforward and can be achieved by reducing the problem to the case with homogeneous boundary conditions and integrating them into the ansatz space, the discretization of the contact conditions is somewhat more intricate. However it should be pointed out, that, compared to standard boundary conditions, the contact conditions actually pose a weaker constraint on the solution, since any $\boldsymbol{\tau}_h$ with vanishing normal trace on Γ_C satisfies the contact conditions. Thus the inf-sup stability (2.21) of our combination remains unaffected, since it holds even if the normal trace on Γ_C is prescribed. We will first write down the abstract problem and then proceed to discuss the treatment of the contact conditions:

Find $\boldsymbol{\sigma}_h \in (\boldsymbol{\Sigma}^C_{\mathbf{g}_N,h})$, $\mathbf{u}_h \in \mathbf{U}_h$, $\boldsymbol{\theta}_h \in \boldsymbol{\Theta}_h$ and $p_h \in P_h$ such that

$$\begin{aligned}(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\mathbf{u}_h, \operatorname{div} \boldsymbol{\tau}_h) + (\boldsymbol{\theta}_h, \mathbf{as}\, \boldsymbol{\tau}_h) + \\ \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, p_h \rangle_{\Gamma_C,h} &= \langle \boldsymbol{\tau}_h \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C}, \\ (\operatorname{div} \boldsymbol{\sigma}_h, \mathbf{v}_h) &= -(\mathbf{f}, \mathbf{v}_h)\ , \\ (\mathbf{as}\, \boldsymbol{\sigma}_h, \boldsymbol{\gamma}_h) &= 0\,, \\ \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} &\le 0\,, \\ \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, p_h \rangle_{\Gamma_C,h} &= 0\,,\end{aligned} \tag{3.13}$$

holds for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}^C_{\mathbf{0},h}$, $\mathbf{v}_h \in \mathbf{U}_h$, $\boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h$ and $w_h \in P_h$.

As for the continuous problem the superscript ^{C} signifies that the frictionless condition (3.1d) is incorporated into the ansatz and test spaces. For an appropriate choice of P_h the KKT system (3.13) is equivalent to the variational inequality (see sections 3.4 and 4.4) that seeks to find $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{K}}^*_{0,h}$ such that

$$(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h - \boldsymbol{\sigma}_h)_\Omega \ge \langle (\boldsymbol{\tau}_h - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle ((\boldsymbol{\tau}_h - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \tag{3.14}$$

holds for all $\boldsymbol{\tau}_h \in \boldsymbol{\mathcal{K}}^*_{0,h}$ with

$$\begin{aligned}\boldsymbol{\mathcal{K}}^*_{0,h} := \{\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}^C_{\mathbf{g}_N,h} : (\operatorname{div} \boldsymbol{\tau}_h, \mathbf{v}_h) &= -(\mathbf{f}, \mathbf{v}_h)\ \ \forall\, \mathbf{v}_h \in \mathbf{U}_h, \\ (\mathbf{as}\, \boldsymbol{\tau}_h, \boldsymbol{\gamma}_h) &= 0\ \forall\, \boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h, \\ \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} &\le 0\ \forall\, w_h \in P_h\}\,.\end{aligned} \tag{3.15}$$

We point out that (3.14) is identical to (3.4) with $\boldsymbol{\mathcal{K}}^*_0$ being replaced by the finite dimensional non-empty, closed, convex set $\boldsymbol{\mathcal{K}}^*_{0,h}$. Thus $\boldsymbol{\sigma}_h$ minimizes $\mathcal{J}^C_{du}$ over $\boldsymbol{\mathcal{K}}^*_{0,h}$ and Theorem 1.12 again guarantees the unique solvability of (3.14).

3.3.1 Discretization of the sign condition

The definition of P_h and of the broken duality product $\langle \cdot, \cdot \rangle_{\Gamma_C,h}$ determine in what sense the sign condition

$$\langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} \le 0 \quad \forall\, w_h \in P_h \tag{3.16}$$

will be enforced in the finite dimensional setting.

The first idea for the choice of P_h would be a conforming discretization, i.e. a subspace of P. In that case $\langle\cdot,\cdot\rangle_{\Gamma_C,h}$ would coincide with the standard duality product $\langle\cdot,\cdot\rangle_{\Gamma_C}$. Because of the positivity requirement in P the only practicable space is $\mathcal{P}_1^+(\mathcal{S}_{h,C})$ since enforcing the positivity for a general element of $\mathcal{P}_k^+(\mathcal{S}_{h,C})$ with $k \geq 2$ is very hard to do numerically. However, since $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n \in \mathcal{DP}_1(\mathcal{S}_{h,C})$ for $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}_h$, testing only against $w_h \in \mathcal{P}_1^+(\mathcal{S}_{h,C})$ leaves room for significant violation of the sign condition which leads to unfeasable numerical solutions. Consequently one has to consider nonconforming options $P_h \not\subset P$ for a stronger enforcement of the contact condition. One can neglect the strong sign requirement of P, or step out of $H^{1/2}(\Gamma_C)$, or both.

An example for relaxing only the sign requirement of the Lagrange multiplier would be setting $P_h = \{w_h \in \mathcal{P}_2(\mathcal{S}_{h,C}) : w_h(\mathbf{z}) \geq 0 \,\forall\, \mathbf{z} \in \Gamma_C\}$ where $\mathbf{z}$ denote again the nodes of the Lagrangian grid (i.e. the vertices and edge midpoints) as in (2.28). In this case the standard duality product would still apply. This choice produces good results in numerical tests for 2D, but it is hard to handle in theory, since it does neither guarantee strong non-negativity of the Lagrange multiplier nor strong non-positivity of the stress approximation. Thus we next consider discretizations beyond $H^{1/2}(\Gamma_C)$.

For $P_h \subset L^2(\Gamma_C)$ the discrete duality product is simply defined as

$$\langle v_h, w_h\rangle_{\Gamma_C,h} := (v_h, w_h)_{L^2(\Gamma_C)} = \sum_{S\in\mathcal{S}_{h,C}} \int_S v_h w_h \, ds\,. \tag{3.17}$$

By testing against positive piecewise affine functions $w_h \in \mathcal{DP}_1^+(\mathcal{S}_{h,C})$ one conserves the strong positivity of the Lagrange multiplier while loosing $H^{1/2}$-regularity. This choice could be considered the most natural one, since it reflects the property conserved by the standard Raviart-Thomas interpolation $\mathcal{R}_h^1$ defined in (1.52).

Lemma 3.3. *Let $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div},\Omega) \cap H^r(\Omega)^{d\times d}$ with $r > \frac{1}{2}$. Let $P_h = \mathcal{DP}_1^+(\mathcal{S}_{h,C})$, then*

$$\langle(\boldsymbol{\tau}\cdot\mathbf{n})_n, w\rangle_{\Gamma_C} \leq 0 \quad \forall\, w \in P \tag{3.18}$$

implies

$$\langle(\mathcal{R}_h^1\boldsymbol{\tau}\cdot\mathbf{n})_n, w_h\rangle_{\Gamma_C,h} \leq 0 \quad \forall\, w_h \in P_h\,. \tag{3.19}$$

Proof. Since $\boldsymbol{\tau} \in \mathbf{H}(\mathrm{div},\Omega) \cap H^r(\Omega)^{d\times d}$ holds, $(\boldsymbol{\tau}\cdot\mathbf{n})_n \in L^2(\Gamma_C)$ and (3.18) implies $(\boldsymbol{\tau}\cdot\mathbf{n})_n \leq 0$ almost everywhere on Γ_C. Thus by definition we have for each $S \in \mathcal{S}_{h,C}$

$$\int_S (\mathcal{R}_h^1\boldsymbol{\tau}\cdot\mathbf{n})_n p_1 \, ds = \int_S (\boldsymbol{\tau}\cdot\mathbf{n})_n p_1 \, ds \leq 0 \tag{3.20}$$

for $p_1 \in \mathcal{P}_1(S)$ with $p_1(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in S$. (3.19) is then a consequence of the definitions in (1.37) and (3.17). □

To better understand what the weak sign condition (3.19) means, we take a look at the situation in 2D on the reference edge $S = [0, 1]$. For $v, w \in \mathcal{P}_1([0, 1])$ we have

$$\int_S vw \, ds = \sum_{i,j=1}^{2} v_i w_j \int_0^1 \psi_i \psi_j \, ds = \frac{1}{6} \sum_{j=1}^{2} w_j \sum_{i=1}^{2} v_i (1 + \delta_{ij}) \tag{3.21}$$

where $\psi_1 = x$ and $\psi_2 = (1 - x)$ represent the nodal basis of $\mathcal{P}_1([0, 1])$. Here $w \geq 0$ on $[0, 1]$ is equivalent to the coefficients being non-negative $w_j \geq 0$, $j = 1, 2$. Thus for non-negative test functions w we have

$$\int_S vw \, ds \leq \frac{1}{6} \max\{2v_1 + v_2, v_1 + 2v_2\} \sum_{j=1}^{2} w_j \quad \forall\, w = \sum_{j=1}^{2} w_j \psi_j \in \mathcal{P}_1^+([0, 1]) \,,$$

implying

$$\int_S vw \, ds \leq 0 \;\; \forall\, w \in \mathcal{P}_1^+([0, 1]) \iff \min_{i=1,2} v_i + 2 \max_{i=1,2} v_i \leq 0 \,, \tag{3.22}$$

which tells us that every violation of the strong sign condition on one end of the edge has to be compensated on the other end of the edge by at least a factor of 2. Applying this observation to the general case tells us that choosing P_h as in Lemma 3.3 allows the sign condition to be violated on up to one third of each edge of Γ_C. The situation in 3D is similar, allowing the sign condition to be violated on up to $\frac{7}{16}$ of each face. We shall see in our numerical experiments, that this is good enough to obtain reasonable results for the frictionless problem but is too weak for the treatment of friction (see chapter 4).

However, (3.22) also suggests that we can obtain the strong sign condition by switching the requirements on v and w. To be more precise we define weakly positive piecewise affine functions on the contact boundary

$$\mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C}) := \{w_h \in \mathcal{DP}_1(\mathcal{S}_{h,C}) : \langle w_h, v_h \rangle_{\Gamma_C,h} \geq 0 \,\forall\, v_h \in \mathcal{DP}_1^+(\mathcal{S}_{h,C})\} \tag{3.23}$$

and state the following Lemma:

Lemma 3.4. *Let $\boldsymbol{\tau}_h$ be in $\boldsymbol{\Sigma}_h$ and let*

$$\langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} \leq 0 \quad \forall\, w_h \in P_h \tag{3.24}$$

hold for $P_h = \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$. Then $\boldsymbol{\tau}_h$ satisfies the strong sign condition (3.1b).

Proof. Since both $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n$ and w_h are defined piecewise, (3.24) can be localized to

$$\int_S (\boldsymbol{\tau}_h \cdot \mathbf{n})_n w_h \, ds \leq 0 \quad \forall\, S \in \mathcal{S}_{h,C} \,,\; \forall\, w_h \in P_h \tag{3.25}$$

and it is enough to show $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n \leq 0$ on an arbitrary side $S \in \mathcal{S}_{h,C}$.

For $d = 2$ S is an edge and we have

$$\int_S (\boldsymbol{\tau}_h \cdot \mathbf{n})_n w_h \, ds = \frac{l_S}{6} (\tau_1 (2w_1 + w_2) + \tau_2 (w_1 + 2w_2)) \tag{3.26}$$

where l_S is the length of the edge S and τ_i, $i = 1, 2$ are the values of $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n$ at the start- and endpoint of S. Since $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n$ is affine on S we only need to verify that $\tau_i \leq 0$ holds for $i = 1, 2$. Suppose the opposite is true, i.e. without loss of generality $\tau_1 > 0$. Setting $w_1 = 2$ and $w_2 = -1$, the arguments leading to (3.22) also yield $w_h \in \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$ and plugging it into (3.26) yields

$$\int_S (\boldsymbol{\tau}_h \cdot \mathbf{n})_n w_h \, ds = \frac{l_S \tau_1}{2} > 0$$

which contradicts (3.25) and thus gives the result.

For $d = 3$ S is a face and we have

$$\int_S (\boldsymbol{\tau}_h \cdot \mathbf{n})_n w_h \, ds = \frac{a_S}{24} \sum_{j=1}^{3} \tau_j \sum_{i=1}^{3} w_i (1 + \delta_{ij})$$

where a_S is the area of the face S. Testing against $w_h = 3\psi_1 - 1\psi_2 - 1\psi_3 \in \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$ gives again the result. □

Remark. For the implementation it is often easier to directly put the constraints on the coefficients of the $\mathcal{RT}$ basis functions than enforcing (3.24). However, Lemma 3.4 provides an interpretation of the enforcing Lagrange multiplier as an element of $\mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$.

3.3.2 A-priori error analysis

In this subsection we will consider to what extend we are able to conserve the approximation properties of the method discussed in chapter 2. Before we tackle the a-priori error analysis we need to introduce an interpolation operator for Raviart-Thomas functions that satisfies the weak symmetry constraint of our finite element combination.

Lemma 3.5. *For $s > 2$ and $k \geq 1$ there exists a bounded interpolation operator $\Pi_h : H(\mathrm{div}, \Omega)^d \cap L^s(\Omega)^{d \times d} \to \mathcal{RT}_k(\mathcal{T}_h)$ with the following properties:*

$$\|\Pi_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{H(\mathrm{div})} \leq C \|\mathcal{R}_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{H(\mathrm{div})} \,, \tag{3.27}$$

$$(\mathbf{as}(\Pi_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}), \boldsymbol{\gamma}_h) = 0 \quad \text{for all } \boldsymbol{\gamma}_h \in \mathcal{P}_k(\mathcal{T}_h)^{d \times d} \cap \boldsymbol{\Theta} \,, \tag{3.28}$$

where $\mathcal{R}_h^k$ is the standard Raviart-Thomas interpolation operator and C is a constant independent of h.

Proof. Let $\tilde{\Pi}_h^k \boldsymbol{\sigma}$ be the minimizer of $\|\boldsymbol{\tau}_h\|_{H(\mathrm{div})}^2$ among all $\boldsymbol{\tau}_h \in \mathcal{RT}_k(\mathcal{T}_h)$ with $\boldsymbol{\tau}_h \cdot \mathbf{n} = 0$ on Γ_N and Γ_C, subject to the following constraints:

$$\begin{aligned} (\mathrm{div}\, \boldsymbol{\tau}_h, \mathbf{v}_h) &= 0 && \text{for all } \mathbf{v}_h \in \mathcal{DP}_k(\mathcal{T}_h) \,, \\ (\mathbf{as}\, \boldsymbol{\tau}_h, \boldsymbol{\gamma}_h) &= (\mathbf{as}\, \boldsymbol{\sigma}, \boldsymbol{\gamma}_h) && \text{for all } \boldsymbol{\gamma}_h \in \mathcal{P}_k(\mathcal{T}_h)^{d \times d} \cap \boldsymbol{\Theta} \,. \end{aligned}$$

The inf-sup stability of the combination guarantees that the solution operator $\tilde{\Pi}_h^k$ is well defined and continuous on $L^2(\Omega)^{d \times d}$. The final operator is then defined as the

composition $\Pi_h^k := \mathcal{R}_h^k + \tilde{\Pi}_h^k \circ (I - \mathcal{R}_h^k)$. The first property (3.27) is then a consequence of the boundedness of $\tilde{\Pi}_h^k$:

$$\begin{aligned}\|\Pi_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{H(\mathrm{div})} &\leq \|\mathcal{R}_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{H(\mathrm{div})} + \|\tilde{\Pi}_h^k(\boldsymbol{\sigma} - \mathcal{R}_h^k \boldsymbol{\sigma})\|_{H(\mathrm{div})} \\ &\leq C\|\mathcal{R}_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{H(\mathrm{div})}\,,\end{aligned}$$

and the second property (3.28) is also inherited:

$$\begin{aligned}(\mathbf{as}(\Pi_h^k \boldsymbol{\sigma} - \boldsymbol{\sigma}), \boldsymbol{\gamma}_h) &= (\mathbf{as}\, \mathcal{R}_h^k \boldsymbol{\sigma} + \mathbf{as}(\tilde{\Pi}_h^k(\boldsymbol{\sigma} - \mathcal{R}_h^k \boldsymbol{\sigma})) - \mathbf{as}\, \boldsymbol{\sigma}, \boldsymbol{\gamma}_h) \\ &= (\mathbf{as}\, \mathcal{R}_h^k \boldsymbol{\sigma} + \mathbf{as}(\boldsymbol{\sigma} - \mathcal{R}_h^k \boldsymbol{\sigma}) - \mathbf{as}\, \boldsymbol{\sigma}, \boldsymbol{\gamma}_h) = 0\end{aligned}$$

for all $\boldsymbol{\gamma}_h \in \mathcal{P}_k(\mathcal{T}_h)^{d\times d} \cap \boldsymbol{\Theta}$. □

We will now generalize Theorem 11.4.2 in [AH09] in order to obtain an a-priori error estimate for the solution of (3.14) in the two-dimensional case.

Theorem 3.6. *Let $d = 2$. Let $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$ be the solutions of (3.4) and (3.14) with either $P_h = \mathcal{DP}_1^+(\mathcal{S}_{h,C})$ or $P_h = \mathcal{P}_1^+(\mathcal{S}_{h,C})$. If the problem is sufficiently regular the following estimate holds with a constant C independent of h and λ:*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\boldsymbol{\Sigma}} \leq C h^{\frac{3}{2}}\,. \tag{3.29}$$

Proof. We will start by defining the residual of the variational inequality as

$$R(\boldsymbol{\tau}, \boldsymbol{\omega}) = (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau} - \boldsymbol{\omega}) - \langle \mathbf{g}_D, (\boldsymbol{\tau} - \boldsymbol{\omega}) \cdot \mathbf{n} \rangle_{\Gamma} - \langle g_C, ((\boldsymbol{\tau} - \boldsymbol{\omega}) \cdot \mathbf{n})_n \rangle_{\Gamma_C} \tag{3.30}$$

where $\boldsymbol{\sigma}$ is the solution of (3.4). Adding up (3.4) and (3.14) yields

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 = (\mathcal{A}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \leq R(\boldsymbol{\tau}, \boldsymbol{\sigma}_h) + R(\boldsymbol{\tau}_h, \boldsymbol{\sigma}) + (\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\tau}_h - \boldsymbol{\sigma})$$

where we can estimate the last term using Cauchy-Schwarz and Young's inequality

$$\begin{aligned}(\mathcal{A}\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\tau}_h - \boldsymbol{\sigma}) &\leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}} \|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathcal{A}} \\ &\leq \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 + \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathcal{A}}^2\end{aligned}$$

to obtain

$$\frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{A}}^2 \leq \inf_{\boldsymbol{\tau} \in \boldsymbol{\mathcal{K}}_0^*} R(\boldsymbol{\tau}, \boldsymbol{\sigma}_h) + \inf_{\boldsymbol{\tau}_h \in \boldsymbol{\mathcal{K}}_{0,h}^*} \left[R(\boldsymbol{\tau}_h, \boldsymbol{\sigma}) + \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\tau}_h\|_{\mathcal{A}}^2 \right]. \tag{3.31}$$

We next use the dual relationships $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{u})$, $\boldsymbol{\theta} = \mathbf{as}\, \boldsymbol{\nabla}\mathbf{u}$ and $p = g_C - u_n$ discussed in the proof of Theorem 3.1 and perform partial integration to rewrite the residual as

$$R(\boldsymbol{\tau}, \boldsymbol{\omega}) = (\boldsymbol{\theta}, \mathbf{as}(\boldsymbol{\omega} - \boldsymbol{\tau})) + \langle p, ((\boldsymbol{\omega} - \boldsymbol{\tau}) \cdot \mathbf{n})_n \rangle_{\Gamma_C}\,, \tag{3.32}$$

which is correct because $\mathrm{div}(\boldsymbol{\omega} - \boldsymbol{\tau}) = \mathbf{0}$ and $(\boldsymbol{\omega} - \boldsymbol{\tau}) \cdot \mathbf{n} = \mathbf{0}$ on Γ_N hold for all $\boldsymbol{\tau}, \boldsymbol{\omega} \in \boldsymbol{\mathcal{K}}_0^* \cup \boldsymbol{\mathcal{K}}_{0,h}^*$. We will verify the result by inserting specific elements of $\boldsymbol{\mathcal{K}}_0^*$ and $\boldsymbol{\mathcal{K}}_{0,h}^*$, namely $\boldsymbol{\sigma}$ and $\Pi_h^1 \boldsymbol{\sigma}$ into (3.31) to obtain upper bounds of the infima.

In order to obtain the desired estimate we will need to handle the following two contact boundary terms:

$$\begin{aligned} cbt_1 :=& \langle p, ((\boldsymbol{\sigma}_h - \boldsymbol{\sigma}) \cdot \mathbf{n})_n \rangle_{\Gamma_C} , \\ cbt_2 :=& \langle p, ((\boldsymbol{\sigma} - \Pi_h^1 \boldsymbol{\sigma}) \cdot \mathbf{n})_n \rangle_{\Gamma_C} , \end{aligned}$$

where Π_h^1 is the interpolation operator defined in Lemma 3.5.

We denote by $\mathcal{I}_h^1$ and π_h^1 the standard nodal interpolation operator and the L^2-orthogonal projection unto piecewise affine functions respectively. We then can estimate cbt_1 in the following way:

$$\begin{aligned} cbt_1 &\leq \langle p - \mathcal{I}_h^1 p, ((\boldsymbol{\sigma}_h - \boldsymbol{\sigma})_n \cdot \mathbf{n}) \rangle_{\Gamma_C} - \langle \mathcal{I}_h^1 p, (\boldsymbol{\sigma} \cdot \mathbf{n})_n \rangle_{\Gamma_C} + \langle \mathcal{I}_h^1 p, (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n \rangle_{\Gamma_C} \\ &\leq \langle \mathcal{I}_h^1 p - p, ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_n \cdot \mathbf{n}) \rangle_{\Gamma_C} + \langle -\mathcal{I}_h^1 p, \pi_h^1 (\boldsymbol{\sigma} \cdot \mathbf{n})_n \rangle_{\Gamma_C} \\ &\leq \| \mathcal{I}_h^1 p - p \|_{1/2} \| (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n} \|_{-1/2} + \langle -\mathcal{I}_h^1 p, \pi_h^1 (\boldsymbol{\sigma} \cdot \mathbf{n})_n \rangle_{\Gamma_C} \\ &\leq C_1 h^{\frac{3}{2}} \| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{\boldsymbol{\Sigma}} + C_2 h^4 , \end{aligned} \tag{3.33}$$

where we used $\mathcal{I}_h^1 p \in P_h \subset \mathcal{DP}_1^+(\mathcal{S}_{h,C})$ for the second inequality and the interpolation estimate (1.48), the trace inequality (1.14) and Lemma 5.2 in [ACS09] for the last inequality. For cbt_2 we obtain

$$cbt_2 \leq \| \mathcal{I}_h^1 p - p \|_{1/2} \| \boldsymbol{\sigma} \cdot \mathbf{n} - \pi_h^1 (\boldsymbol{\sigma} \cdot \mathbf{n}) \|_{-1/2} \leq C_3 h^4 \tag{3.34}$$

using the fact that $\Pi_h^1 \boldsymbol{\sigma} \cdot \mathbf{n} = \mathcal{R}_h^1 \boldsymbol{\sigma} \cdot \mathbf{n} = \pi_h^1 (\boldsymbol{\sigma} \cdot \mathbf{n})$ on Γ and the interpolation estimates (1.48) and (1.49).

We now can estimate the first term in (3.31) using (3.32) and applying the Cauchy-Schwarz-inequality:

$$\begin{aligned} \inf_{\boldsymbol{\tau} \in \boldsymbol{\mathcal{K}}_0^*} R(\boldsymbol{\tau}, \boldsymbol{\sigma}_h) &\leq (\boldsymbol{\theta}, \mathbf{as}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma})) + cbt_1 \leq (\boldsymbol{\theta} - \mathcal{I}_h^1 \boldsymbol{\theta}, \mathbf{as}(\boldsymbol{\sigma}_h - \boldsymbol{\sigma})) + cbt_1 \\ &\leq \| \boldsymbol{\theta} - \mathcal{I}_h^1 \boldsymbol{\theta} \|_0, \| \boldsymbol{\sigma}_h - \boldsymbol{\sigma} \|_0 + cbt_1 \\ &\leq C_4 h^2 \| \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \|_{\boldsymbol{\Sigma}} + cbt_1 . \end{aligned} \tag{3.35}$$

Because of Lemma 3.3 and Lemma 3.5 we may insert $\Pi_h^1 \boldsymbol{\sigma} \in \boldsymbol{\mathcal{K}}_{0,h}^*$ for $\boldsymbol{\tau}_h$ into the second term in (3.31) and obtain:

$$R(\Pi_h^1 \boldsymbol{\sigma}, \boldsymbol{\sigma}) \leq \inf_{\boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h} (\boldsymbol{\theta} - \boldsymbol{\gamma}_h, \mathbf{as}(\boldsymbol{\sigma} - \Pi_h^1 \boldsymbol{\sigma})) + cbt_2 \tag{3.36}$$

$$\leq C_5 h^2 \| \boldsymbol{\sigma} - \mathcal{R}_h^1 \boldsymbol{\sigma} \|_{\boldsymbol{\Sigma}} \leq C_6 h^4 + cbt_2 . \tag{3.37}$$

where we used (1.50). Finally we recall that (2.18) and the continuity of the bilinear form $(\mathcal{A} \cdot, \cdot)$ imply the equivalence of the energy norm $\| \cdot \|_{\mathcal{A}}$ to $\| \cdot \|_{\boldsymbol{\Sigma}} = \| \cdot \|_{H(\mathrm{div})}$ on $\{ \boldsymbol{\tau} \in \boldsymbol{\Sigma}_{\mathbf{0}} : \operatorname{div} \boldsymbol{\tau} = \mathbf{0} \}$. Thus, putting together (3.33)-(3.37) and completing the square yields the result. □

Theorem 3.6 yields only a suboptimal convergence rate and doesn't cover the choice of $P_h = \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$. However, in the case that $\boldsymbol{\sigma}_h$ happens to satisfy the strong sign condition (3.1b), $cbt_1 \leq 0$ holds and we obtain the optimal estimate $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\boldsymbol{\Sigma}} \leq Ch^2$. This also means that if it is possible to construct an interpolation operator Π_h^1 that additionally to respecting the divergence and weak symmetry constraints also conserves the strong sign condition while retaining optimal approximation qualities, one could use the same strategy to prove an optimal estimate for the choice of $P_h = \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$. As of yet, we have not succeeded in constructing such an operator.

Nevertheless, it is important to point out that the regularity requirements on the solution necessary to obtain (3.29) (i.e. $\boldsymbol{\sigma} \in H^2(\Omega)$) are almost never met (cf. [Rös00]). Hence the need to use adaptive refinement procedures and appropriate a-posteriori error estimators. We will see in our numerical tests in section 3.5 that such procedures enable us to recover the optimal rate for both $P_h = \mathcal{DP}_1^{+}(\mathcal{S}_{h,C})$ and $P_h = \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$, while the choice of $P_h = \mathcal{P}_1^{+}(\mathcal{S}_{h,C})$ doesn't produce feasible results at all, despite being covered by Theorem 3.6.

3.3.3 A-posteriori error analysis

In this section we will discuss the modifications necessary to generalize the result in Theorem 2.4 to frictionless contact problems. As already alluded to in chapter 2 the main difference is that the boundary term in (2.33) does in general no longer vanish, due to the fact that the contact conditions on Γ_C provide significantly weaker information on the behaviour of the solution on the boundary. Thus, our goal will be to bound it by computable quantities in order to ensure the reliabilty of the proposed error estimator.

Theorem 3.7. *Let $\boldsymbol{\sigma}$ solve (3.5) and $\boldsymbol{\sigma}_h$ solve (3.13). Let $\boldsymbol{\sigma}_h$ satisfy (3.1b) and $\mathbf{u}_h^R \in (\mathbf{u}_D + \mathbf{K}_{g_C})$. Then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathrm{div})} \leq C\left(\eta_1 + \eta_2 + \eta_3 + \eta_4\right) \tag{3.38}$$

holds with η_1, η_2, η_3 as in Theorem 2.4 and

$$\eta_4 := \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C \rangle_{\Gamma_C}^{1/2},$$

and a constant C which is independent of λ and h.

Proof. We examine the boundary term in (2.33) using $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_{\mathbf{0}}^C$ and $\mathbf{u} - \mathbf{u}_h^R \in \mathbf{V}$ to restrict it to Γ_C.

$$\begin{aligned}
\langle (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{u} - \mathbf{u}_h^R \rangle_{\Gamma} &= \langle (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{u} - \mathbf{u}_h^R \rangle_{\Gamma_C} \\
&= \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_n, (\mathbf{u} - \mathbf{u}_h^R)_n \rangle_{\Gamma_C} \\
&= \langle ((\boldsymbol{\sigma}_h - \boldsymbol{\sigma}) \cdot \mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C + g_C - u_n \rangle_{\Gamma_C} \\
&= \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C \rangle_{\Gamma_C} + \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} \\
&\quad + \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_n, g_C - (\mathbf{u}_h^R)_n \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C}.
\end{aligned}$$

The last term vanishes due to the last equation in (3.5) and the second and third term are non positive since both $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$ satisfy the strong sign condition. Thus we have

$$\langle(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n}, \mathbf{u}-\mathbf{u}_h^R\rangle_\Gamma \leq \langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C\rangle_{\Gamma_C} \tag{3.39}$$

and the result follows by the arguments used in the proof of Theorem 2.4. □

Remark. All the modifications discussed in the remarks after Theorem 2.4 apply also to Theorem 3.7.

Since

$$\langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C\rangle_{\Gamma_C} = \langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C\rangle_{\Gamma_{C,h}}$$

holds, the localisation of η_4 is straightforward:

$$\eta_{4,T} = ((\boldsymbol{\sigma}_h\cdot\mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C)^{1/2}_{L^2(S_T)} \quad \text{with } S_T = \partial T \cap \Gamma_C. \tag{3.40}$$

While we can adopt most of the displacement reconstruction algorithm discussed in section 2.4.1, the condition $\mathbf{u}_h^R \in (\mathbf{u}_D + \mathbf{K}_{g_C})$ requires a slight modification of step *(iii)*. In addition to enforcing the boundary conditions on Γ_D we also need to ensure $g_C - (\mathbf{u}_h^R)_n \geq 0$ on Γ_C. If we additionally enforce the condition $(\mathbf{u}_h^R)_n = g_C$ on S whenever $(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n \neq 0$ on S for all $S \in \mathcal{S}_{h,C}$ we obtain $\eta_4 = 0$ and we are left with the same error estimator as for the problem of linear elasticity.

3.4 Solution method

In this section we will turn to the question of how to obtain a solution of (3.14). Simple projection based fixed point iterations are sometimes used to solve variational inequalities. Besides being rather slow, this approach is not suitable for our problem since a projection onto the admissible set would be quite intricate due to the divergence and symmetry constraints. Multilevel approaches for the displacement based and Least Squares formulations of contact problems have been proposed in [Kra09, Kra06, RKSK18]. Since an extension of these methods to the dual formulation is beyond the scope of this work, we chose to adopt the approach proposed in [HIK02] to transform (3.13) into a system of non linear equations that can be solved using a Newton type method.

Before we will recapitulate the ideas in [HIK02] and discuss their application to our problem, we introduce some notation that will condense the equations in (3.13) enabling us to focus on the method in Euclidean space. Furthermore, we will assume $\mathbf{g}_N = \mathbf{0}$ in order to simplify notation and refer to the modification described in (2.11) for the general case.

Let $\{\boldsymbol{\phi}_i\}_{i=1}^{n_\Sigma}$ be a basis of $\boldsymbol{\Sigma}^C_{\mathbf{0},h}$, $\{\boldsymbol{\psi}_i\}_{i=1}^{n_U}$ a basis of $\mathbf{U}_h$ and $\{\boldsymbol{\chi}_i\}_{i=1}^{n_\Theta}$ a basis of $\boldsymbol{\Theta}_h$. We can then contract the L^2 products of (3.13) into matrix vector multiplications using the following notation:

$$\begin{aligned} \boldsymbol{\sigma}_h &= \sum_{j=1}^{n_\Sigma} x_j\boldsymbol{\phi}_j & \mathbf{u}_h &= \sum_{j=1}^{n_U} y_j\boldsymbol{\psi}_j & \boldsymbol{\theta}_h &= \sum_{j=1}^{n_\Theta} z_j\boldsymbol{\chi}_j \\ A_{ij} &= (\mathcal{A}\boldsymbol{\phi}_j, \boldsymbol{\phi}_i) & B_{ij} &= (\operatorname{div}\boldsymbol{\phi}_j, \boldsymbol{\psi}_i) & C_{ij} &= (\mathbf{as}\,\boldsymbol{\phi}_j, \boldsymbol{\chi}_i) \end{aligned} \tag{3.41}$$

where j runs from 1 to n_Σ and i runs from 1 to the respective dimensions n_Σ, n_U and n_Θ. We can also express the sign condition by a matrix vector relation depending on the choice of P_h. For all choices of P_h discussed in section 3.3.1 except $\mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$ we can easily characterize P_h as the positive linear span of the nodal basis of the underlying polynomial space. Let for example $\{\rho_i\}_{i=1}^m$ be the nodal basis of $\mathcal{DP}_1(\mathcal{S}_{h,C})$. Then we have

$$\mathcal{DP}_1^{+}(\mathcal{S}_{h,C}) = \left\{ w_h = \sum_{i=1}^{m} w_i \rho_i : w_i \geq 0,\ i = 1, ..., m \right\}$$

and

$$Dx \leq 0 \iff \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} \leq 0 \quad \forall\, w_h \in P_h \tag{3.42}$$

with

$$D_{ij} = \langle (\boldsymbol{\phi}_j \cdot \mathbf{n})_n, \rho_i \rangle_{\Gamma_C,h} \quad i = 1, ..., m \quad j = 1, ..., n_\Sigma\,. \tag{3.43}$$

Vector inequalities such as $Dx \leq 0$ are here and for the rest of this work to be understood in terms of components. Analogous to (3.41), p_h can be written as

$$p_h = \sum_{i=1}^{m} p_i \rho_i\,. \tag{3.44}$$

In the case of $P_h = \mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C})$ a similar characterization is possible:

$$\mathcal{DP}_1^{\oplus}(\mathcal{S}_{h,C}) = \left\{ w_h = \sum_{i=1}^{m} w_i \rho_i : Mw_i \geq 0,\ i = 1, ..., m \right\},$$

with

$$M_{ij} = \langle \rho_j, \rho_i \rangle_{\Gamma_C,h} \quad i, j = 1, ..., m\,. \tag{3.45}$$

However, if one uses a basis of $\mathbf{\Sigma}_{\mathbf{0},h}^{C}$ for which a mapping $k : j \mapsto i$ exists, that satisfies

$$(\boldsymbol{\phi}_j \cdot \mathbf{n})_n\big|_S \equiv \begin{cases} \rho_{k(j)} & \text{if } S \cap \operatorname{supp}(\boldsymbol{\phi}_j) \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3.46}$$

for all $S \in \mathcal{S}_{h,C}$, the m inequality constraints can be decoupled from each other and the strong sign condition can be enforced directly using the following definiton of D

$$D_{ij} = \begin{cases} 1 & \text{if } i = k(j) \\ 0 & \text{otherwise} \end{cases} \quad i = 1, ..., m \quad j = 1, ..., n_\Sigma \tag{3.47}$$

instead of (3.43). In this case

$$p_h = \sum_{i=1}^{m} (M^{-1}p)_i \rho_i \tag{3.48}$$

holds.

The standard $\mathcal{RT}_1(\mathcal{T}_h)^d$-basis can easily be modified to satisfy (3.46) by local coordinate transformations (cf. [Kra06]). Finally we summarize the right hand side by setting

$$\begin{aligned} b_i &= \langle \boldsymbol{\phi}_i \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle (\boldsymbol{\phi}_i \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C}, \\ f_i &= -(\mathbf{f}, \boldsymbol{\psi}_i) \end{aligned}$$

and can now write (3.13) as

$$Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p = b\,, \tag{3.49a}$$
$$Bx = f\,, \tag{3.49b}$$
$$Cx = 0\,, \tag{3.49c}$$
$$Dx \leq 0\,, \tag{3.49d}$$
$$p \geq 0\,, \tag{3.49e}$$
$$p^\mathsf{T} Dx = 0\,. \tag{3.49f}$$

The existence of a unique solution $\xi = (x, y, z, p) \in \mathbb{R}^n = (\mathbb{R}^{n_\Sigma} \times \mathbb{R}^{n_U} \times \mathbb{R}^{n_\Theta} \times \mathbb{R}^m)$ of (3.49) can be inferred from the unique solvability of (3.14) using results from convex optimization which will be discussed in section 4.4. The first step towards the application of the semismooth Newton Method suggested in [HIK02] is the following observation :

Lemma 3.8. *(3.49e)-(3.49f) can equivalently be expressed as*

$$p - \max\{0, p + \kappa(Dx)\} = 0 \tag{3.50}$$

for any fixed $\kappa > 0$. The max*-operation is understood componentwise.*

Proof. Both (3.50) and (3.49e)-(3.49f) imply $(Dx)_i = 0$ or $p_i = 0$ for each $i \in \{1, ..., m\}$. Whenever $(Dx)_i = 0$ holds, $0 = p_i - \max\{0, p_i + \kappa(Dx)_i\} = p_i - \max\{0, p_i\}$ is equivalent to $p_i \geq 0$. Whenever $p_i = 0$ holds, $0 = p_i - \max\{0, p_i + \kappa(Dx)_i\} = -\max\{0, \kappa(Dx)_i\}$ is equivalent to $\kappa(Dx)_i \leq 0$ which is equivalent to $(Dx)_i \leq 0$. □

Remark. While the size of κ evidently doesn't change the solution of (3.49) it can have an effect on the performance of the semismooth Newton method which we will be introduced in the next section. Whenever the constraint $D(x) = Dx$ is linear, κ can in fact be chosen completely arbitrarily. However, whenever the constraint is nonlinear (as will be the case for problems with friction) the active set (see (3.60)) depends on the choice of κ. Thus it should be used to balance the magnitudes of p and $D(x)$.

As a consequence of Lemma 3.8 the KKT system (3.49) is equivalent to

$$Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p - b = 0\,, \tag{3.51a}$$
$$Bx - f = 0\,, \tag{3.51b}$$
$$Cx = 0\,, \tag{3.51c}$$
$$p - \max\{0, p + \kappa(Dx)\} = 0\,, \tag{3.51d}$$

which can be summarized into

$$F(\xi) = 0 \tag{3.52}$$

with the non-linear function

$$F : \mathbb{R}^n \to \mathbb{R}^n \quad \xi \mapsto F(\xi) := \begin{pmatrix} Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p - b \\ Bx - f \\ Cx \\ p - \max\{0, p + \kappa(Dx)\} \end{pmatrix} . \tag{3.53}$$

3.4.1 Semismooth Newton Method

Since the function $s \mapsto \max\{0, s\}$ is clearly not continuously differentiable, (3.52) doesn't fit into the framework of the standard Newton method. Therefore, a generalized concept of differentiation is needed.

Definition 3.9. A mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ is called *Newton-differentiable* in the open subset $U \subset \mathbb{R}^n$ if there exists a mapping $DF : U \to \mathbb{R}^{m \times n}$ such that

$$NL_F(\xi) := \lim_{\eta \to 0} \frac{\|F(\xi + \eta) - F(\xi) - DF(\xi + \eta)\eta\|}{\|\eta\|} = 0 \tag{3.54}$$

for every $\xi \in U$. DF is called *Newton-derivative* .

Example 3.10. The mapping

$$DF(s) = \begin{cases} 1, & s > 0 \\ 0, & s \leq 0 \end{cases} \tag{3.55}$$

is a Newton-derivative of the function $F(s) := \max\{0, s\}$. Indeed, since F is smooth everywhere else we only need to consider the point $s = 0$:

$$NL_F(0) = \begin{cases} \lim\limits_{\eta \searrow 0} \frac{\|F(\eta) - DF(\eta)\eta\|}{\|\eta\|} = \lim\limits_{\eta \searrow 0} \frac{\|\eta - \eta\|}{\|\eta\|} = 0 \\ \lim\limits_{\eta \nearrow 0} \frac{\|F(\eta) - DF(\eta)\eta\|}{\|\eta\|} = \lim\limits_{\eta \nearrow 0} \frac{\|0 - 0\|}{\|\eta\|} = 0 \end{cases} . \tag{3.56}$$

Example 3.11. The mapping

$$DF(s) = \begin{cases} \frac{s}{F(s)}, & s \neq 0 \\ 0, & s = 0 \end{cases} \tag{3.57}$$

is a Newton-derivative of the Euclidean norm $F(s) := \sqrt{\sum_{i=1}^n s_i^2}$:

$$NL_F(0) = \lim_{\eta \to 0} \frac{\|F(\eta) - DF(\eta)\eta\|}{\|\eta\|} = \lim_{\eta \to 0} \frac{\| |\eta| - \frac{|\eta|^2}{|\eta|} \|}{\|\eta\|} = 0 . \tag{3.58}$$

We point out that DF is not necessarily unique and that whenever F is continuously differentiable on U, F' is a Newton-derivative of F. Furthermore the Newton-derivative is clearly linear and for locally Lipschitz continuous functions the chain rule holds:

Lemma 3.12. *Let both $G : \mathbb{R}^l \to \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ be Newton-differentiable in $\mathbb{R}^l$, $\mathbb{R}^n$ respectively. Let furthermore G be locally Lipschitz continuous and DF be locally bounded. Then the composition $H := F \circ G : \mathbb{R}^l \to \mathbb{R}^m$ is also Newton-differentiable on $\mathbb{R}^l$ with Newton-derivative $DH = (DF \circ G)DG$.*

Proof. For arbitrary $\xi \in \mathbb{R}^l$ and $\eta \in B_\varepsilon(0)$ (the open ball around 0 with radius ε) we define $x := G(\xi)$ and $y = y(\eta) := G(\xi + \eta) - G(\xi)$. Thus we can write

$$\begin{aligned}
&\|H(\xi+\eta) - H(\xi) \\
&\quad -DH(\xi+\eta)\eta\| = \|F(G(\xi+\eta)) - F(G(\xi)) \\
&\qquad\qquad - DF(G(\xi+\eta))DG(\xi+\eta)\eta\| \\
&\qquad\qquad = \|F(x+y) - F(x) - DF(x+y)y \\
&\qquad\qquad + DF(x+y)y - DF(G(\xi+\eta))DG(\xi+\eta)\eta\| \\
&\qquad\qquad = \|F(x+y) - F(x) - DF(x+y)y \\
&\qquad\qquad + DF(G(\xi+\eta))[G(\xi+\eta) - G(\xi) - DG(\xi+\eta)\eta]\| \\
&\qquad\qquad \le \|F(x+y) - F(x) - DF(x+y)y\| \\
&\qquad\qquad + \|DF(G(\xi+\eta))\|\|G(\xi+\eta) - G(\xi) - DG(\xi+\eta)\eta\|
\end{aligned}$$

and, using the Lipschitz continuity of G on $\overline{B_\varepsilon(\xi)}$, we obtain

$$\begin{aligned}
NL_H(\xi) &= \lim_{\eta\to 0} \frac{\|H(\xi+\eta) - H(\xi) - DH(\xi+\eta)\eta\|}{\|\eta\|} \\
&\le \lim_{\eta\to 0} \frac{\|F(x+y(\eta)) - F(x) - DF(x+y(\eta))y(\eta)\|}{\|y(\eta)\|} \frac{\|y(\eta)\|}{\|\eta\|} \\
&\quad + \lim_{\eta\to 0} \frac{\|DF(G(\xi+\eta))\|\|G(\xi+\eta) - G(\xi) - DG(\xi+\eta)\eta\|}{\|\eta\|} \\
&\le L \lim_{y\to 0} \frac{\|F(x+y) - F(x) - DF(x+y)y\|}{\|y\|} \\
&\quad + C \lim_{\eta\to 0} \frac{\|G(\xi+\eta) - G(\xi) - DG(\xi+\eta)\eta\|}{\|\eta\|} = 0
\end{aligned}$$

where L is the local Lipschitz constant of G and

$$C = \sup_{x \in G\left(\overline{B_\varepsilon(\xi)}\right)} \|DF(x)\|,$$

which is finite because DF is locally bounded. □

The following theorem justifies the use of Newton's method to find a solution of (3.52).

Theorem 3.13. *Suppose that ξ_* is a solution to $F(\xi) = 0$ and that F is Newton-differentiable in an open neighborhood U containing ξ_* with Newton-derivative DF. If $DF(\xi)$ is nonsingular for all $\xi \in U$ and $\{DF(\xi)^{-1} : \xi \in U\}$ is bounded, then the Newton iteration*

$$DF(\xi_k)\delta\xi_k = -F(\xi_k)\,, \tag{3.59a}$$
$$\xi_{k+1} = \xi_k + \delta\xi_k \tag{3.59b}$$

converges superlinearly to ξ_*, provided that $\|\xi_0 - \xi_*\|$ is sufficiently small.

Proof. See [HIK02]. □

In order to apply (3.13) we need to find a Newton-derivative of (3.53). For the first three slots the standard derivative works, while for the last slot we can use example 3.10 to build a Newton-derivative. For this purpose we define the set of active indices

$$S_\xi := \{i \in \{1, ..., m\} : p_i + \kappa(Dx)_i > 0\} \tag{3.60}$$

and its complement

$$S_\xi^c := \{i \in \{1, ..., m\} : i \notin S_\xi\} \tag{3.61}$$

and introduce the following notation: For the identity matrix $I \in \mathbb{R}^{m\times m}$ the subscript S or S^c inidcates that all rows, except those corresponding to the indices in S or S^c respectively, are set to zero. For $m = 3$ and $S = \{1, 3\}$ this would mean

$$I_S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad I_{S^c} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.62}$$

With this we can write the Newton-derivative of (3.53):

$$DF(\xi) = \begin{pmatrix} A & B^\mathsf{T} & C^\mathsf{T} & D^\mathsf{T} \\ B & 0 & 0 & 0 \\ C & 0 & 0 & 0 \\ -\kappa I_{S_\xi} D & 0 & 0 & I_{S_\xi^c} \end{pmatrix}. \tag{3.63}$$

Under the assumption that D is surjective (which is clearly the case for $P_h \subset \mathcal{DP}_1(\mathcal{S}_{h,C})$), the regularity of $DF(\xi)$ can be inferred from (2.21) and (2.18). Indeed, if $\eta^\circ = (x^\circ, y^\circ, z^\circ, p^\circ)$ solves the homogeneous equation $DF(\xi)\eta^\circ = 0$ the last slot immediately yields

$$\left.\begin{aligned} (Dx^\circ)_i = 0 \quad i \in S_\xi \\ p_i^\circ = 0 \quad i \in S_\xi^c \end{aligned}\right\} \implies x^{\circ\mathsf{T}} D^\mathsf{T} p^\circ = 0\,.$$

Together with slots 1, 2 and 3 this implies

$$0 = x^{\circ\mathsf{T}}(Ax^\circ + B^\mathsf{T} y^\circ + C^\mathsf{T} z^\circ + D^\mathsf{T} p^\circ) = x^{\circ\mathsf{T}} A x^\circ\,. \tag{3.64}$$

(2.18) implies that A is positive definite on $\ker B$ and thus (3.64) implies $x^\circ = 0$. Let N be a matrix whose columns are a basis of $\ker D$, then the first slot yields

$$N^\mathsf{T} B^\mathsf{T} y^\circ + N^\mathsf{T} C^\mathsf{T} z^\circ = 0\,. \tag{3.65}$$

The argumentation at the beginning of section 3.3 together with (2.21) implies the surjectivity of

$$\begin{pmatrix} BN \\ CN \end{pmatrix}$$

which together with (3.65) implies $y^\circ = 0$ and $z^\circ = 0$. The assumption on D implies $p^\circ = 0$ and finally $\eta^\circ = 0$. Thus $DF(\xi)$ is invertible for all $\xi \in \mathbb{R}^n$ and Theorem 3.13 yields the superlinear convergence of Newton's method for "good" initial values.

The structure of (3.63) allows the interpretation of the semismooth Newton method as an active set algorithm: Each Newton iterate ξ_{k+1} solves a standard problem of linear elasticity with homogeneous boundary conditions (possibly only weakly enforced, dependent on the choice of P_h) on the part of Γ_C represented by the active set $S_k = S_{\xi_k}$. This observation also suggests the use of $\|p - \max\{0, p + \kappa(Dx)\}\| \leq \varepsilon_{tol}$ as a stopping criterion since the first three equations in (3.51) will be satisfied exactly in each step. In our computations we observed that as long as $D(x)$ is at least piecewise linear (e.g. for the frictional problem in 2D) the exact solution with $p - \max\{0, p + \kappa(Dx)\} = 0$ is usually obtained after a justifiable number of iterations.

3.5 Numerical Experiments

In this section we report some results achieved by our implementation of the discussed finite element method. As in chapter 2 a Dörfler marking strategy is used for the adaptive refinement procedure.

Example 1: Hertzian Contact - half-disk on straight line

In this first example we consider the domain Ω of the lower half-disk with center at the origin and radius $R = 0.5$. The body is constrained by a rigid foundation represented by the horizontal line at $x_2 = -0.5$ and the potential contact boundary is $\Gamma_C := \left\{(R\cos(\varphi), R\sin(\varphi)) : \varphi \in \left(\frac{4}{3}\pi, \frac{5}{3}\pi\right)\right\}$. While surface traction forces $\mathbf{g}_N$ on $\Gamma_N := \Gamma \setminus \left(\overline{\Gamma_D \cup \Gamma_C}\right)$ as well as volume forces $\mathbf{f}$ on Ω are set to zero, displacement on $\Gamma_D := (-R, R) \times \{0\}$ is prescribed by $\mathbf{g}_D = (0, -0.01)^\mathsf{T}$. Again both the compressible case ($\lambda = 1$) and the incompressible case ($\lambda = \infty$) will be treated with the shear modulus μ being scaled to 1. Furthermore we perfomed computations both with $P_h = \mathcal{DP}_1^+$ and $P_h = \mathcal{DP}_1^\oplus$ in order to compare the effects of the weak and strong imposition of the sign condition. Again a Dörfler parameter of $\theta = 0.8$ was used in all computations.

Table 3.1 shows the results of the compressible case for $P_h = \mathcal{DP}_1^\oplus(\mathcal{S}_{h,C})$. Table 3.2 reports the incompressible case. The number of active constraints for the sign condition are denoted by A_n. Since the displacement was always reconstructed in a way such that η_4 vanished, we did not list it in the Table (cf. end of section 3.3.3).

Figure 3.3 depicts the initial and deformed configurations of the incompressible case after 5 steps of adaptive refinement. The deformed configuration is obtained using the reconstructed displacement $\mathbf{u}_h^R \in \mathcal{P}_2(\mathcal{T}_h)$. Besides the contact zone refinement concentrates at the corner singularities. In the contact zone the refinement concentrates at the transition point from contact to separation as can be seen in Figure3.2.

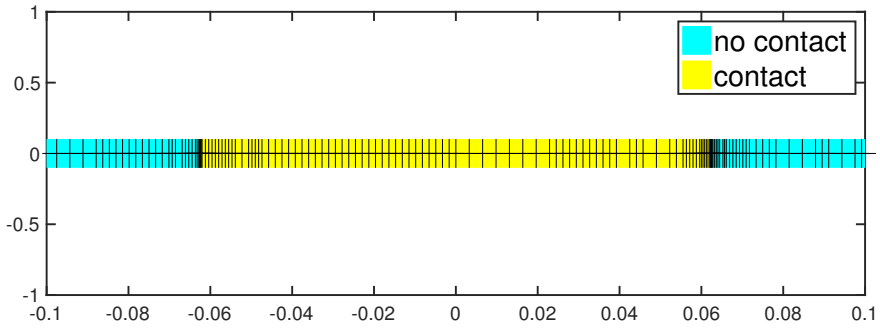

Figure 3.2: Example 1: Contact zone after after 11 refinements for $\lambda = \infty$.

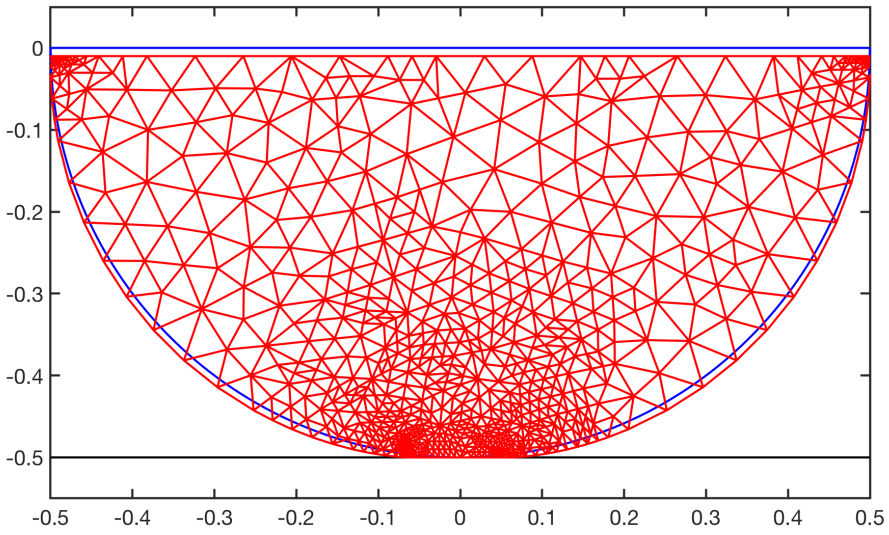

Figure 3.3: Example 1: Deformed mesh after 5 refinements for $\lambda = \infty$.

The surface forces in the highly resolved contact zone as well as the corresponding reconstructed displacements are shown in figures 3.5 and 3.6. The expected elliptic distribution of the contact pressure is recovered in both cases, with significantly higher peak pressure in the incompressible case. However, the contact zone is slightly broader in the incompressible case, which does not correspond directly to the formulas in [KO88, Chap-

ter 6] taken from [Gol60, Chapter IV]. We strongly suspect this to be due to the different boundary conditions (displacement vs. forces prescribed on top). Furthermore a significant qualitative difference of the tangential displacements between the compressible and incompressible case can be observed: While for $\lambda = 1$ the tangential displacement increases even faster in the contact zone, it almost vanishes for $\lambda = \infty$.

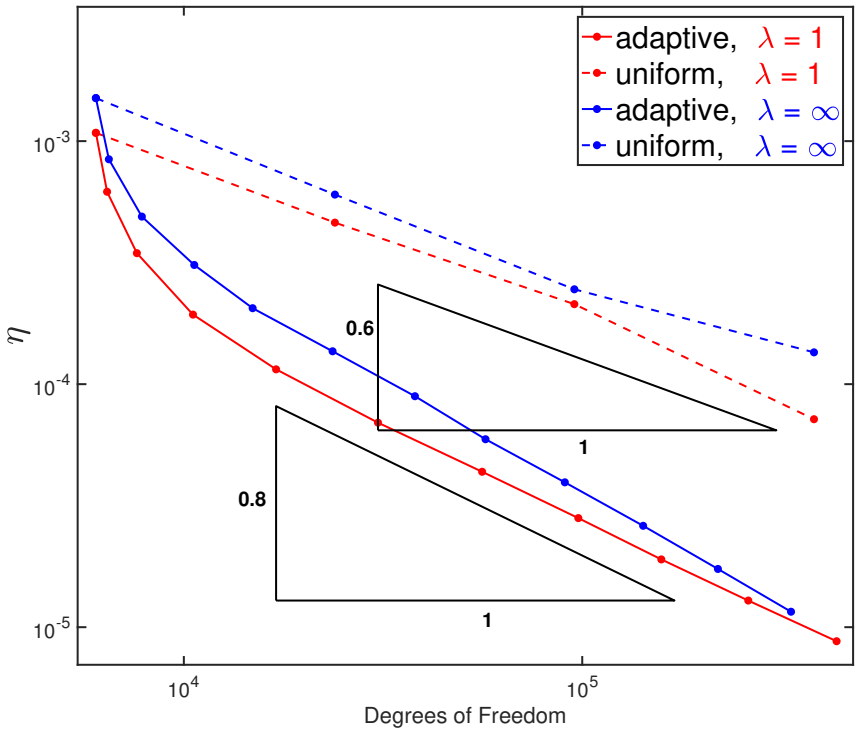

Figure 3.4: Example 1: Adaptive vs. uniform refinement.

A comparison of the reduction of η for uniform and adaptive refinement is given in Figure3.4. Even though the convergence behaviour for adaptive refinement is clearly better than for uniform refinement, the optimal convergence behaviour achievable $\eta \sim N_h^{-1}$ (for N_h being the number of degrees of freedom) is not quite reached in the compressible case. As the second example suggests, this is probably due to the curved boundary being resolved only by a piecewise linear curve. For optimal convergence rates it would be necessary to use a piecewise quadratic approximation of the boundary and corresponding parametric Raviart-Thomas elements for the stress approximation (cf. [BMS14]).

The results for $P_h = \mathcal{DP}_1^+(\mathcal{S}_{h,C})$ are summarized in 3.3 and 3.4. The comparison in Figure3.7 shows that, while the method with $P_h = \mathcal{DP}_1^+(\mathcal{S}_{h,C})$ has some trouble in the beginning, comparable rates are obtained once the mesh is fine enough. However, with this choice of P_h, strictly speaking, η is not a reliable upper bound for the error $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|$, since $\boldsymbol{\sigma}_h$ does not necessarily satisfy the assumptions of Theorem 3.7, due to possible violations of the strong sign condition as illustrated in Figure3.8.

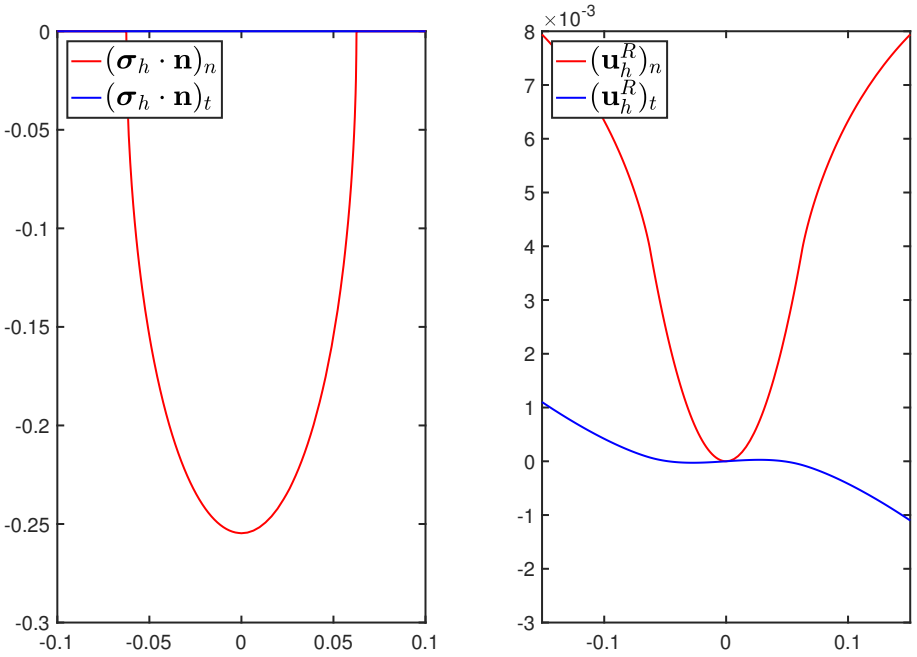

Figure 3.5: Example 1: Stress and displacement in contact zone for $\lambda = \infty$.

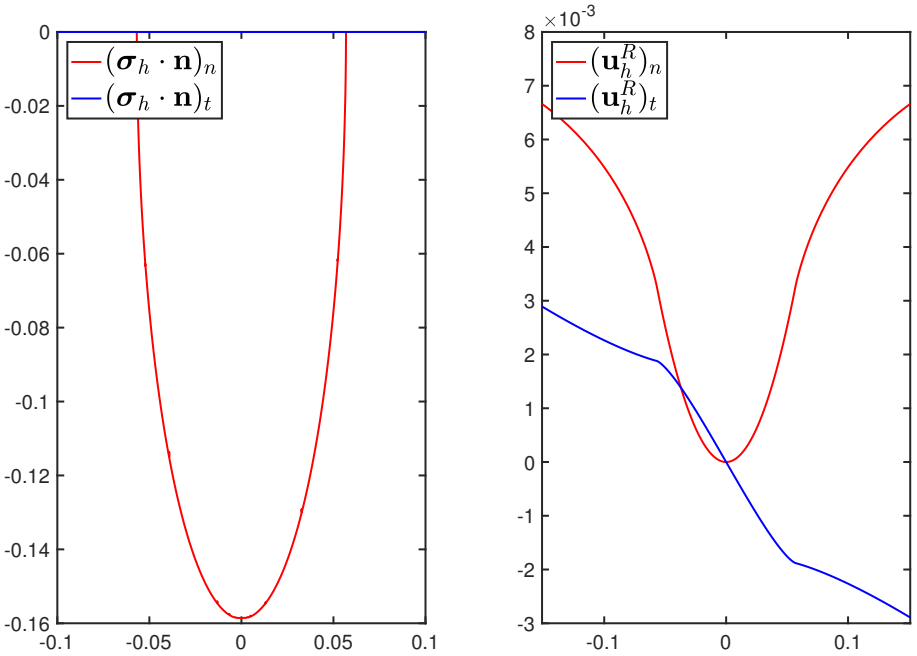

Figure 3.6: Example 1: Stress and displacement in contact zone for $\lambda = 1$.

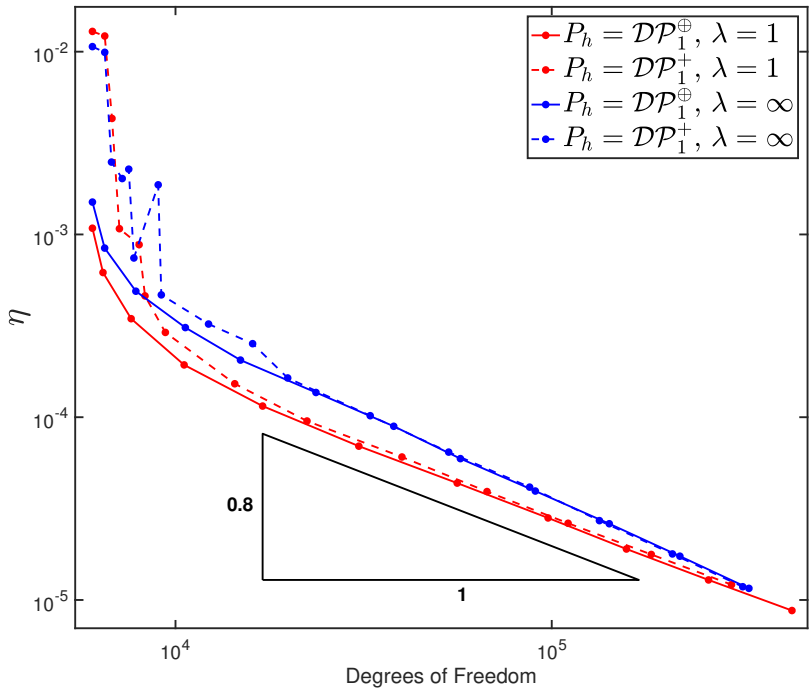

Figure 3.7: Example 1: Weak vs. strong enforcement of sign condition

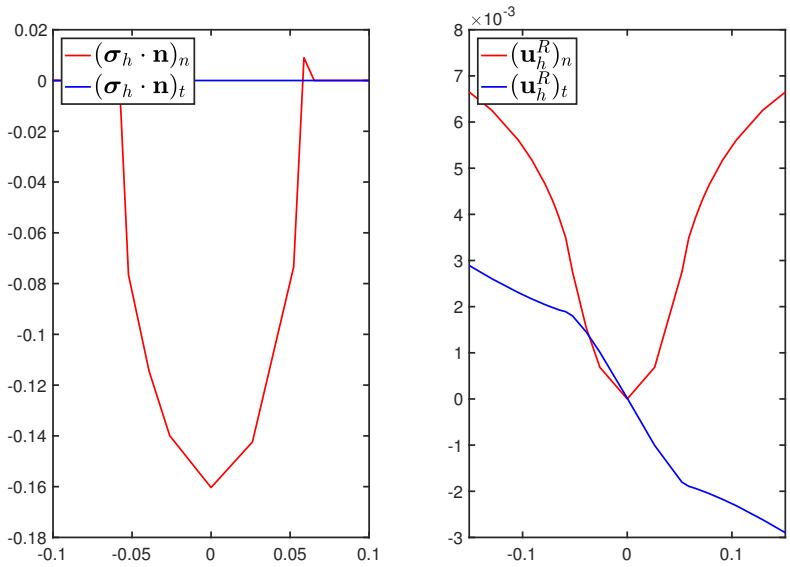

Figure 3.8: Example 1: Violation of strong sign condition for $P_h = \mathcal{DP}_1^{+}$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	η_1	η_2	η_3	η	eoc
0	361	3604	2166	205	16	0.000800422	0.000500822	0.000527716	0.00108166	
1	384	3834	2304	219	22	0.000432408	0.00031731	0.000308885	0.000618928	8.63
2	456	4554	2736	257	24	0.000240723	0.000191184	0.000158729	0.000345968	3.38
3	630	6294	3780	353	34	0.000124315	0.000105915	0.000103368	0.00019328	1.8
4	1020	10194	6120	557	42	7.43857e-05	6.46348e-05	5.94252e-05	0.000115075	1.08
5	1840	18394	11040	985	58	4.46427e-05	3.91992e-05	3.59419e-05	6.9436e-05	0.86
6	3360	33612	20160	1784	84	2.79244e-05	2.49451e-05	2.2353e-05	4.36084e-05	0.77
7	5860	58608	35160	3098	118	1.81162e-05	1.62952e-05	1.40117e-05	2.8108e-05	0.79
8	9456	94558	56736	5000	149	1.21918e-05	1.0984e-05	9.58212e-06	1.90028e-05	0.82
9	15608	156090	93648	8270	175	8.27332e-06	7.47415e-06	6.43042e-06	1.28709e-05	0.78
10	25954	259536	155724	13821	247	5.63469e-06	5.12762e-06	4.31167e-06	8.75402e-06	0.76

Table 3.1: Example 1: Results for compressible case ($\lambda = 1$)

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	η_1	η_2	η_3	η	eoc
0	361	3604	2166	205	16	0.00110568	0.000773289	0.000663198	0.00150344	
1	388	3874	2328	221	20	0.000594634	0.000440305	0.00040208	0.000842097	7.74
2	469	4684	2814	268	24	0.0003169	0.000246465	0.000279412	0.000489123	2.85
3	635	6346	3810	357	32	0.000203909	0.000163831	0.000165136	0.000309337	1.51
4	890	8902	5340	493	42	0.000133643	0.000108698	0.000111823	0.000205378	1.21
5	1416	14168	8496	766	56	8.86396e-05	7.28656e-05	7.40286e-05	0.000136553	0.88
6	2283	22840	13698	1223	76	5.77549e-05	4.81267e-05	4.80154e-05	8.92036e-05	0.89
7	3439	34404	20634	1818	86	3.84918e-05	3.137e-05	3.25025e-05	5.93473e-05	1
8	5450	54504	32700	2844	112	2.54997e-05	2.10634e-05	2.15078e-05	3.94524e-05	0.89
9	8577	85766	51462	4453	146	1.68676e-05	1.40833e-05	1.41407e-05	2.61307e-05	0.91
10	13227	132260	79362	6816	176	1.12088e-05	9.25786e-06	9.49175e-06	1.7362e-05	0.95
11	20197	201950	121182	10351	224	7.45367e-06	6.19896e-06	6.3289e-06	1.15775e-05	0.96

Table 3.2: Example 1: Results for incompressible case ($\lambda = \infty$)

l	$\lvert\mathcal{T}_h\rvert$	A_n	DOF	η	eoc
0	361	14	6015	0.0129083	
1	389	18	6486	0.0121794	0.77
2	406	22	6780	0.00431265	23.42
3	424	30	7095	0.001075	30.59
4	478	35	8004	0.000879141	1.67
5	495	38	8298	0.000461035	17.89
6	560	42	9399	0.000291028	3.69
7	858	52	14371	0.000152425	1.52
8	1337	60	22365	9.53262e-05	1.06
9	2395	72	39985	6.05818e-05	0.78
10	4031	96	67299	3.91889e-05	0.84
11	6616	124	110528	2.62864e-05	0.8
12	10998	154	183894	1.77446e-05	0.77
13	17948	198	300460	1.21276e-05	0.78

Table 3.3: Example 1: Results for $P_h = \mathcal{DP}_1^+$ and $\lambda = 1$

l	$\lvert\mathcal{T}_h\rvert$	A_n	DOF	η	eoc
0	361	14	6015	0.0106582	
1	389	18	6486	0.00992045	0.95
2	405	22	6759	0.00248839	33.54
3	432	27	7218	0.00202206	3.16
4	449	29	7512	0.00227548	-2.96
5	463	34	7752	0.000743338	35.57
6	537	33	8995	0.00186839	-6.2
7	547	38	9169	0.000467072	72.36
8	731	46	12241	0.000323442	1.27
9	958	54	16032	0.000252558	0.92
10	1188	65	19872	0.000163954	2.01
11	1974	76	32925	0.000101921	0.94
12	3199	96	53215	6.44278e-05	0.96
13	5261	118	87340	4.144e-05	0.89
14	8073	149	133891	2.71731e-05	0.99
15	12636	184	209306	1.78609e-05	0.94
16	19445	220	321814	1.18151e-05	0.96
17	29290	264	484505	7.88765e-06	0.99

Table 3.4: Example 1: Results for $P_h = \mathcal{DP}_1^+$ and $\lambda = \infty$

Example 2: Rectangle on circular arc

Our next example is basically the situation from the first one flipped upside down. The body of interest is now a rectangle with length $2R$ and width R while the rigid foundation takes the shape of a circular arc with radius R. Material parameters as well as Dirichlet data are the same as in example 1. Initial and deformed configurations of the compressible case after 5 refinement steps are depicted in Figure3.9, and the refinement of the contact zone is given in Figure3.10.

Figure 3.11 presents again a comparison of the reduction of η for uniform and adaptive refinement, where now the optimal rate $\eta \sim N_h^{-1}$ is achieved, illustrating the efficiency of the proposed error estimator. The numerical results are summarized in Table 3.5, where DOF represents the sum of all degrees of freedom.

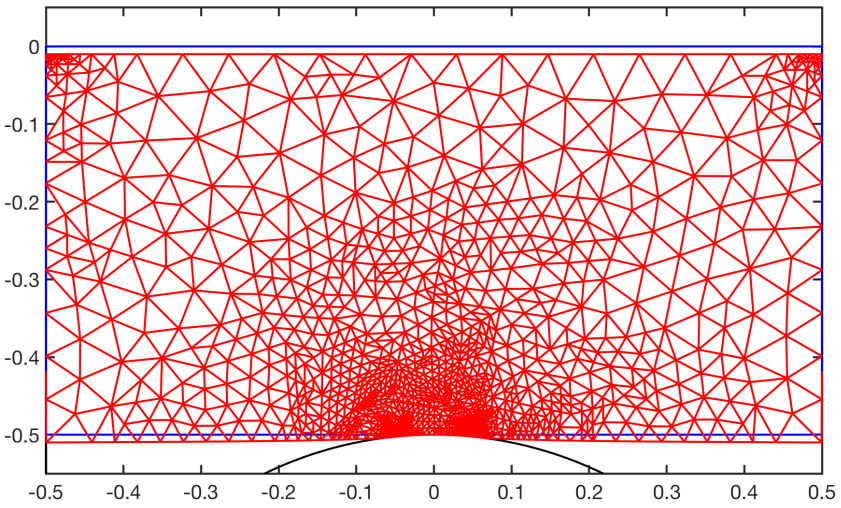

Figure 3.9: Example 2: Deformed mesh after 5 refinements for $\lambda = 1$.

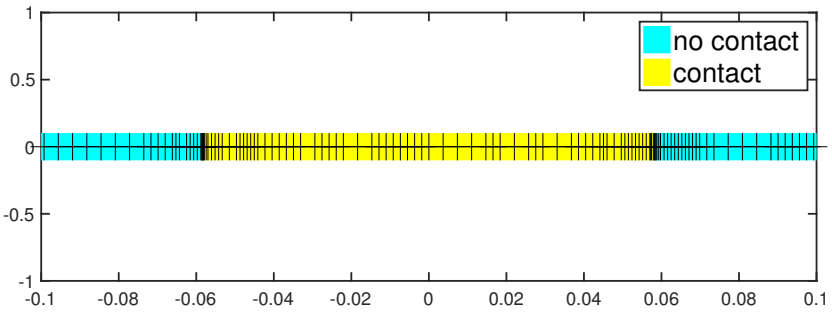

Figure 3.10: Example 2: Contact zone after 9 refinements for $\lambda = 1$.

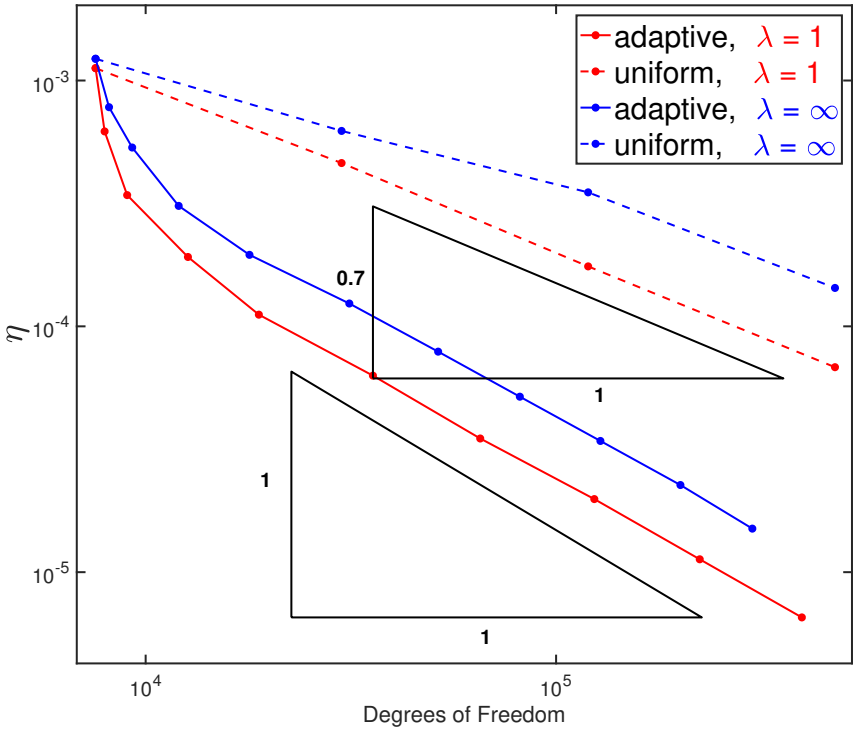

Figure 3.11: Example 2: Adaptive vs. uniform refinement.

l	$\lvert\mathcal{T}_h\rvert$	A_n	DOF	η	eoc
0	452	30	7551	0.00112262	
1	475	32	7944	0.000620793	11.68
2	538	40	9015	0.000341888	4.72
3	757	54	12678	0.000191492	1.7
4	1130	62	18873	0.000111557	1.36
5	2148	100	35789	6.30151e-05	0.89
6	3933	124	65338	3.49742e-05	0.98
7	7470	194	123914	1.9829e-05	0.89
8	13514	252	223867	1.12812e-05	0.95
9	23991	322	396984	6.545e-06	0.95

Table 3.5: Example 2: Results for compressible case ($\lambda = 1$)

Chapter 4

Towards Coulomb-Friction

In this chapter the frictionless condition (3.1d) will be replaced by the Coulomb friction law. This law models the relationship between contact pressure and frictional shear stress, and the resulting restriction of displacements tangential to the contact surface. We refer the reader to [KO88] for a detailed discussion of this and alternative friction models. The classical formulation of Coulomb's friction law is summarized by the following conditions:

$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| \le \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n| \qquad \text{on } \Gamma_C\,, \tag{4.1a}$$
$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| < \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n| \Rightarrow \mathbf{u}_t = 0\,, \tag{4.1b}$$
$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| = \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n| \Rightarrow \exists\, c \ge 0 : \mathbf{u}_t = -c(\boldsymbol{\sigma}\cdot\mathbf{n})_t\,, \tag{4.1c}$$

where μ_F is the material- and surface-dependent *coefficient of friction* and $|\cdot|$ represents the Euclidean norm or the absolute value depending on the dimension of the argument.

The first condition (4.1a) ensures that the maximal static friction force is proportional to the contact pressure and we will subsequently refer to it as the *friction condition*. It is thus clear that large friction coefficients correspond to a rough surface while small friction coefficients correspond to a smooth surface with the frictionless problem of chapter 3 as the limit case. (4.1b) states that before the critical value is reached there is no tangential displacement - the solid *"sticks"* to the obstacle - and finally (4.1c) desribes *"sliding"* into the opposite direction of the shear stress whenever it is large enough. The conditions in (4.1) can be equivalently expressed in a more manageable form:

$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| \le \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n| \qquad \text{on } \Gamma_C\,, \tag{4.2a}$$
$$(\boldsymbol{\sigma}\cdot\mathbf{n})_t\cdot\mathbf{u}_t = -\mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n||\mathbf{u}_t| \quad \text{on } \Gamma_C\,, \tag{4.2b}$$

and we will refer to (4.2b) as *frictional complementarity condition*. Combining (4.2) with the equations in (3.2) we obtain the *classical problem with Coulomb friction*:

Find sufficiently smooth $\mathbf{u}$ and $\boldsymbol{\sigma}$ such that

$$\begin{aligned}
\operatorname{div}\boldsymbol{\sigma} &= -\mathbf{f} && \text{in } \Omega,\\
\mathbf{as}\,\boldsymbol{\sigma} &= \mathbf{0} && \text{in } \Omega,\\
\mathcal{A}\boldsymbol{\sigma} &= \boldsymbol{\varepsilon}(\mathbf{u}) && \text{in } \Omega,\\
\boldsymbol{\sigma}\cdot\mathbf{n} &= \mathbf{g}_N && \text{on } \Gamma_N,\\
\mathbf{u} &= \mathbf{g}_D && \text{on } \Gamma_D,\\
g_C - \mathbf{u}_n &\geq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n &\leq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n(g_C - \mathbf{u}_n) &= 0 && \text{on } \Gamma_C,\\
|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| - \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n| &\leq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_t\cdot\mathbf{u}_t + \mu_F|(\boldsymbol{\sigma}\cdot\mathbf{n})_n||\mathbf{u}_t| &= 0 && \text{on } \Gamma_C.
\end{aligned} \tag{4.3}$$

Coulomb friction confronts us with two major challenges. First, besides not being continuously differentiable, the absolute value in the friction condition (4.1a) has no apparent meaning for $\boldsymbol{\sigma}\cdot\mathbf{n} \in \mathbf{H}^{-1/2}(\Gamma_C)$. We will see that these issues can be circumvented in the weak formulations, such that they do not impair the theory in the continuous setting and we can extend most results from the previous chapters. However, the resulting non-smooth nonlinearity in the discrete setting will require appropriate numerical treatment and obstructs the derivation of a-priori estimates.

Second, the nature of the friction condition is such, that both the primal and the dual variational formulations require the a-priori knowledge of the contact pressure $(\boldsymbol{\sigma}\cdot\mathbf{n})_n$ of the solution, thus creating a bootstrap. We therefore introduce the auxiliary *problem with given friction* for a given contact pressure g_F:

Find sufficiently smooth $\mathbf{u}$ and $\boldsymbol{\sigma}$ such that

$$\begin{aligned}
\operatorname{div}\boldsymbol{\sigma} &= -\mathbf{f} && \text{in } \Omega,\\
\mathbf{as}\,\boldsymbol{\sigma} &= \mathbf{0} && \text{in } \Omega,\\
\mathcal{A}\boldsymbol{\sigma} &= \boldsymbol{\varepsilon}(\mathbf{u}) && \text{in } \Omega,\\
\boldsymbol{\sigma}\cdot\mathbf{n} &= \mathbf{g}_N && \text{on } \Gamma_N,\\
\mathbf{u} &= \mathbf{g}_D && \text{on } \Gamma_D,\\
g_C - \mathbf{u}_n &\geq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n &\leq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_n(g_C - \mathbf{u}_n) &= 0 && \text{on } \Gamma_C,\\
|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| - \mu_F|g_F| &\leq 0 && \text{on } \Gamma_C,\\
(\boldsymbol{\sigma}\cdot\mathbf{n})_t\cdot\mathbf{u}_t + \mu_F|g_F||\mathbf{u}_t| &= 0 && \text{on } \Gamma_C.
\end{aligned} \tag{4.4}$$

The problem with given friction is sometimes called Tresca problem or problem with Tresca friction (see [DL76, Cap14]). We will now proceed to analyse (4.4) and subsequently use it to define weak solutions of (4.3).

4.1 Variational formulations with given friction

As mentioned above, the friction condition

$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| - \mu_F|g_F| \leq 0 \text{ on } \Gamma_C \tag{4.5}$$

has no evident meaning for $\boldsymbol{\sigma}\in\mathbf{H}(\mathrm{div},\Omega)$ and $g_F\in H^{-1/2}(\Gamma_C)$. However, if we assume that g_F has a sign, more specifically if $g_F\in P'$, then we simply have $|g_F| = -g_F$. Since the restriction of the normal trace of $\boldsymbol{\sigma}\in\mathbf{H}(\mathrm{div},\Omega)$ to the contact boundary is well defined in our setting (see chapter 1), the following *weak friction condition*

$$\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{r}\rangle_{\Gamma_C} + \langle\mu_F g_F,|\mathbf{r}|\rangle_{\Gamma_C} \leq 0 \quad \forall\,\mathbf{r}\in\mathbf{Q}\,, \tag{4.6}$$

is well defined and equivalent to (4.5) whenever $(\boldsymbol{\sigma}\cdot\mathbf{n})_t\in\mathbf{L}^2(\Gamma_C)$ and $g_F\in L^2(\Gamma_C)$.

Indeed, if $(\boldsymbol{\sigma}\cdot\mathbf{n})_t\in\mathbf{L}^2(\Gamma_C)$ and $g_F\in L^2(\Gamma_C)\cap P'$ holds, then $|(\boldsymbol{\sigma}\cdot\mathbf{n})_t|\in L^2(\Gamma_C)$ and $\mu_F g_F = -\mu_F|g_F|\in L^2(\Gamma_C)$. (4.5) can then be interpreted in an "almost everywhere" sense and immediately implies (4.6). Conversely, if there exists a subset $\Gamma_+\subset\Gamma_C$ of positive measure with

$$|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| - \mu_F|g_F| > 0 \text{ on } \Gamma_+$$

then

$$\begin{aligned}((\boldsymbol{\sigma}\cdot\mathbf{n})_t,(\boldsymbol{\sigma}\cdot\mathbf{n})_t)_{\Gamma_+} - (\mu_F|g_F|,|(\boldsymbol{\sigma}\cdot\mathbf{n})_t|)_{\Gamma_+}\\ = (|(\boldsymbol{\sigma}\cdot\mathbf{n})_t| - \mu_F|g_F|,|(\boldsymbol{\sigma}\cdot\mathbf{n})_t|)_{\Gamma_+}\\ = a_* > 0\end{aligned}$$

holds. Due to the density of $\mathbf{C}_c^\infty(\Gamma_+)$ in $\mathbf{L}^2(\Gamma_+)$ there exists a sequence $\{\mathbf{r}_k\}\subset\mathbf{C}_c^\infty(\Gamma_+)\subset\mathbf{Q}$ with $\mathbf{r}_k\to(\boldsymbol{\sigma}\cdot\mathbf{n})_t$ in $\mathbf{L}^2(\Gamma_+)$ and consequently also $|\mathbf{r}_k|\to|(\boldsymbol{\sigma}\cdot\mathbf{n})_t|$ in $L^2(\Gamma_+)$. Defining

$$a_k := \langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{r}_k\rangle_{\Gamma_C} - \langle\mu_F|g_F|,|\mathbf{r}_k|\rangle_{\Gamma_C} = ((\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{r}_k)_{\Gamma_+} - (\mu_F|g_F|,|\mathbf{r}_k|)_{\Gamma_+}$$

we thus obtain $a_k\to a_*$. This implies the existence of k_0 with $a_{k_0}>0$ and consequently the existence of $\mathbf{r}_{k_0}\in\mathbf{Q}$ with

$$\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{r}_{k_0}\rangle_{\Gamma_C} - \langle\mu_F|g_F|,|\mathbf{r}_{k_0}|\rangle_{\Gamma_C} > 0\,.$$

Thus (4.6) is a meaningful generalization of (4.5) which can be included in the admissible set for the dual problem.

Stress-based formulation

The *dual problem with given friction* consists in finding $\boldsymbol{\sigma}\in\boldsymbol{\mathcal{K}}^*_{g_F}$ such that

$$\mathcal{J}^C_{du}(\boldsymbol{\sigma}) = \min_{\boldsymbol{\tau}\in\boldsymbol{\mathcal{K}}^*_{g_F}}\mathcal{J}^C_{du}(\boldsymbol{\tau}) := \frac{1}{2}(\mathcal{A}\boldsymbol{\tau},\boldsymbol{\tau})_\Omega - \langle\boldsymbol{\tau}\cdot\mathbf{n},\mathbf{g}_D\rangle_\Gamma - \langle(\boldsymbol{\tau}\cdot\mathbf{n})_n,g_C\rangle_{\Gamma_C}\,. \tag{4.7}$$

Compared to (3.3) the only thing that changed is the admissible set (see (1.25)). Consequently also the variational inequality remains the same:

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau} - \boldsymbol{\sigma})_\Omega \geq \langle (\boldsymbol{\tau} - \boldsymbol{\sigma}) \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle ((\boldsymbol{\tau} - \boldsymbol{\sigma}) \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\mathcal{K}}^*_{g_F} \,. \tag{4.8}$$

In the KKT conditions, however, the friction constraint is enforced by an additional Lagrange multiplier $\mathbf{q} \in \mathbf{Q}$:

Find $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_N + \hat{\boldsymbol{\sigma}}) \in (\boldsymbol{\sigma}_N + \boldsymbol{\Sigma}^F_{\mathbf{0}})$, $\mathbf{u} \in \mathbf{U}$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $p \in P$, and $\mathbf{q} \in \mathbf{Q}$ such that

$$\begin{aligned}
&(\mathcal{A}\hat{\boldsymbol{\sigma}}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div} \boldsymbol{\tau}) + (\boldsymbol{\theta}, \mathbf{as}\, \boldsymbol{\tau}) + \\
&\quad \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_t, \mathbf{q} \rangle_{\Gamma_C} = \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} - (\mathcal{A}\boldsymbol{\sigma}_N, \boldsymbol{\tau}) \,, \\
&\qquad\qquad (\operatorname{div} \hat{\boldsymbol{\sigma}}, \mathbf{v}) = -\left(\mathbf{f} + \operatorname{div} \boldsymbol{\sigma}_N, \mathbf{v}\right) \,, \\
&\qquad\qquad (\mathbf{as}\, \hat{\boldsymbol{\sigma}}, \boldsymbol{\gamma}) = -\left(\mathbf{as}\, \boldsymbol{\sigma}_N, \boldsymbol{\gamma}\right) \,, \\
&\qquad\qquad \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, w \rangle_{\Gamma_C} \leq 0 \,, \\
&\qquad\qquad \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} = 0 \,, \\
&\qquad \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_t, \mathbf{r} \rangle_{\Gamma_C} + \langle \mu_F g_F, |\mathbf{r}| \rangle_{\Gamma_C} \leq 0 \,, \\
&\qquad \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_t, \mathbf{q} \rangle_{\Gamma_C} + \langle \mu_F g_F, |\mathbf{q}| \rangle_{\Gamma_C} = 0
\end{aligned} \tag{4.9}$$

holds for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}^F_{\mathbf{0}}$, $\mathbf{v} \in \mathbf{U}$, $\boldsymbol{\gamma} \in \boldsymbol{\Theta}$, $w \in P$ and $\mathbf{r} \in \mathbf{Q}$.

Displacement formulation

Let again $\mathbf{u}_D \in \mathbf{V}_{\mathbf{g}_D}$. Then the *primal problem with given friction* consists in finding $\mathbf{u} \in (\mathbf{u}_D + \mathbf{K}_{g_C})$ such that

$$\mathcal{J}^F_{pr}(\mathbf{u}) = \min_{\mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C})} \mathcal{J}^F_{pr}(\mathbf{v}) := \mathcal{J}_{pr}(\mathbf{v}) - \langle \mu_F g_F, |\mathbf{v}_t| \rangle_{\Gamma_C} \,. \tag{4.10}$$

Here the admissible set remained the same as in the frictionless case (3.6), but the energy functional as well as the variational inequality feature now a nonsmooth term representing the frictional energy:

$$(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega - \langle \mu_F g_F, |\mathbf{v}_t| - |\mathbf{u}_t| \rangle_{\Gamma_C} \geq (\mathbf{f}, \mathbf{v} - \mathbf{u})_\Omega + (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} \,, \tag{4.11}$$

for all $\mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C})$.

4.1.1 Duality

Just like for the frictionless case a strong duality result holds:

Theorem 4.1.

(i) Let $\mathbf{u}$ *be the solution of (4.11), then* $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u})$, $\mathbf{u}$, $\boldsymbol{\theta} := \mathbf{as}\, \boldsymbol{\nabla} \mathbf{u}$, $p := g_C - u_n$ *and* $\mathbf{q} := -\mathbf{u}_t$ *solve (4.9).*

(ii) Let conversely $\boldsymbol{\sigma}$, $\mathbf{u}$, $\boldsymbol{\theta}$, p *and* $\mathbf{q}$ *be the solution of (4.9), then* $\mathbf{u}$ *is in* $H^1(\Omega)$ *and solves (4.11).*

Proof. We restrict the proof again to the case $\lambda < \infty$ since the treatment of the incompressible case is analogous to the proof of remark after (2.1).

(i) Let $\mathbf{u}$ be the solution of (4.11). Then starting as in the proof of Theorem 3.1 we obtain $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}^{\mathbf{f}} \cap \boldsymbol{\Sigma}^F_{\mathbf{g}_N}$ and

$$\langle \hat{\boldsymbol{\sigma}} \cdot \mathbf{n}, \hat{\mathbf{v}} - \hat{\mathbf{u}} \rangle_{\Gamma_C} - \langle \mu_F g_F, |\hat{\mathbf{v}}_t| - |\hat{\mathbf{u}}_t| \rangle_{\Gamma_C} \geq 0 \quad \forall\, \hat{\mathbf{v}} \in \mathbf{H}^{1/2}(\Gamma_C) \text{ s.t. } g_C - \hat{v}_n \in P \tag{4.12}$$

with $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_N$ and $\hat{\mathbf{u}} := \mathbf{u} - \mathbf{u}_D$. Restriction to the normal components again yields

$$\begin{aligned} \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, w \rangle_{\Gamma_C} &\leq 0 \quad \forall\, w \in P\,, \\ \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, g_C - u_n \rangle_{\Gamma_C} &= 0\,. \end{aligned}$$

Finally testing in (4.12) with $\hat{\mathbf{v}} = \hat{u}_n \mathbf{n}$ and $\hat{\mathbf{v}} = \hat{\mathbf{u}} + \hat{\mathbf{u}}_t$ we obtain

$$- \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_t, \hat{\mathbf{u}}_t \rangle_{\Gamma_C} + \langle \mu_F g_F, |\hat{\mathbf{u}}_t| \rangle_{\Gamma_C} = 0 \tag{4.13}$$

and thus

$$\langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_t, \hat{\mathbf{v}} \rangle_{\Gamma_C} - \langle \mu_F g_F, |\hat{\mathbf{v}}_t| \rangle_{\Gamma_C} \geq 0 \quad \forall\, \hat{\mathbf{v}} \in \mathbf{Q} \tag{4.14}$$

which implies the last two relations in (4.9) by setting $\mathbf{q} = -\hat{\mathbf{u}}_t$ and $\mathbf{r} = -\hat{\mathbf{v}}$.

What remains to be shown is the first equation in (4.9). Proceeding as in the proof of Theorem 3.1 leads to

$$\begin{aligned} (\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}) + (\mathbf{u}, \operatorname{div} \boldsymbol{\tau}) + (\boldsymbol{\theta}, \operatorname{as} \boldsymbol{\tau}) + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_t, \mathbf{q} \rangle_{\Gamma_C} = \\ \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{u}_D \rangle_{\Gamma} + \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{u} \rangle_{\Gamma_C} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C - u_n \rangle_{\Gamma_C} - \langle (\boldsymbol{\tau} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} = \\ \langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{g}_D \rangle_{\Gamma} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \end{aligned}$$

for all $\boldsymbol{\tau} \in \boldsymbol{\Sigma}^F_{\mathbf{0}}$.

(ii) Let now $\boldsymbol{\sigma}$, $\mathbf{u}$, $\boldsymbol{\theta}$, p and $\mathbf{q}$ solve (4.9). By testing with appropriate test functions in the first equation of (4.9) and applying partial integration we again obtain $\boldsymbol{\nabla} \mathbf{u} = \mathcal{A}\boldsymbol{\sigma} + \boldsymbol{\theta}$, $\hat{\mathbf{u}} = \mathbf{u} - \mathbf{u}_D \in V$ and

$$\langle \boldsymbol{\tau} \cdot \mathbf{n}, \mathbf{u} \rangle_{\Gamma_C} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} + \langle (\boldsymbol{\tau} \cdot \mathbf{n})_t, \mathbf{q} \rangle_{\Gamma_C} = \langle (\boldsymbol{\tau} \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \quad \forall\, \boldsymbol{\tau} \in \boldsymbol{\Sigma}^C_{\mathbf{0}}\,. \tag{4.15}$$

Testing only with $\boldsymbol{\tau} \in C^\infty(\Omega)^{d \times d}$ such that $(\boldsymbol{\tau} \cdot \mathbf{n})_n$ vanishes on Γ_C yields $\mathbf{u}_t = \hat{\mathbf{u}}_t = -\mathbf{q}$ on Γ_C. Analogously, we obtain again $g_C - \hat{u}_n = p$ on Γ_C and hence $\mathbf{u} - \mathbf{u}_D \in \mathbf{K}_{g_C}$. Finally, proceeding again as in the proof of Theorem 3.1, we get

$$\begin{aligned} (\mathbf{f}, \mathbf{v} - \mathbf{u})_\Omega &= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} - \langle \hat{\boldsymbol{\sigma}} \cdot \mathbf{n}, \hat{\mathbf{v}} - \hat{\mathbf{u}} \rangle_{\Gamma_C} \\ &= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} + \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_t, (-\hat{\mathbf{v}}_t) - \mathbf{q} \rangle_{\Gamma_C} \\ &\quad + \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, g_C - \hat{v}_n \rangle_{\Gamma_C} - \langle (\hat{\boldsymbol{\sigma}} \cdot \mathbf{n})_n, p \rangle_{\Gamma_C} \\ &\leq (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} - \langle \mu_F g_F, |\mathbf{v}_t| - |\mathbf{q}| \rangle_{\Gamma_C} \\ &= (\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v} - \mathbf{u}))_\Omega - (\mathbf{g}_N, \mathbf{v} - \mathbf{u})_{\Gamma_N} - \langle \mu_F g_F, |\mathbf{v}_t| - |\mathbf{u}_t| \rangle_{\Gamma_C} \end{aligned}$$

for all $\mathbf{v} = \mathbf{u}_D + \hat{\mathbf{v}}$ mit $\hat{\mathbf{v}} \in \mathbf{K}_{g_C}$, where for the inequality we used $(-\hat{\mathbf{v}}_t) \in \mathbf{Q}$ and the last four relations in (3.5). □

4.1.2 Existence and Uniqueness

While the existence and uniqueness of the solution of (4.7) follows with the same arguments as for the frictionless problem, the situation for the primal problem (4.10) requires additional arguments to obtain the weak lower semicontinuity of the non linear term $\langle \mu_F g_F, |\mathbf{v}_t| \rangle_{\Gamma_C}$ that allows an application of Theorem 1.11. We refer the reader to [HHN96, KO88] and remark that again the well posedness of the primal problem implies the unique existence of the Lagrange multipliers in (4.9) via Theorem 4.1. We again summarize these observations in the following theorem:

Theorem 4.2. *Both the dual problem with given friction (4.7) or equivalently (4.9) and the primal problem with given friction (4.10) or equivalently (4.11) admit a unique solution.*

4.2 Variational formulations with Coulomb friction

Using the well posed problem with given friction as a tool, we can now proceed to precisely formulate what is regarded as a solution to the problem with Coulomb friction, following the reasoning used in [HHN96]. To do this, we observe that for otherwise fixed data both the solution of the primal and dual problems with given friction admit the interpretation as mappings $\mathcal{S}_{pr}$ and $\mathcal{S}_{du}$ that map a functional $g_F \in P'$ to the respective solutions in the following sense:

$$\mathcal{S}_{pr} : \begin{cases} P' \to \mathbf{H}^1(\Omega) \\ g_F \mapsto \mathcal{S}_{pr}(g_F) = \mathbf{u} \end{cases} \tag{4.16}$$

and

$$\mathcal{S}_{du} : \begin{cases} P' \to \mathbf{H}(\mathrm{div}, \Omega) \\ g_F \mapsto \mathcal{S}_{du}(g_F) = \boldsymbol{\sigma} \end{cases} \tag{4.17}$$

We point out that Theorem 4.1 tells us that

$$\mathcal{C}\boldsymbol{\varepsilon}(\mathcal{S}_{pr}(g_F)) = \mathcal{S}_{du}(g_F) \tag{4.18}$$

holds. Thus we can consider another mapping that relates the given slip stress g_F to the resulting contact pressure of the solution:

$$\Psi : \begin{cases} P' \to P' \\ g_F \mapsto \Psi(g_F) = (\mathcal{C}\boldsymbol{\varepsilon}(\mathcal{S}_{pr}(g_F)) \cdot \mathbf{n})_n = (\boldsymbol{\sigma} \cdot \mathbf{n})_n \end{cases} \tag{4.19}$$

Definition 4.3. Let g_F^* be a solution of the equation

$$\mu_F \Psi(g_F) = \mu_F g_F \tag{4.20}$$

then $\mathcal{S}_{pr}(g_F^*)$ is called a solution of the *primal problem with Coulomb friction* and $\mathcal{S}_{du}(g_F^*)$ is called a solution of the *dual problem with Coulomb friction.*

The primal problem with Coulomb friction can formally be expressed in terms of a quasi-variational inequality:

Find $\mathbf{u} \in (\mathbf{u}_D + \mathbf{K}_{g_C})$ such that

$$(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}), \boldsymbol{\varepsilon}(\mathbf{v}-\mathbf{u}))_\Omega - \langle \mu_F(\mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u})\cdot\mathbf{n})_n, |\mathbf{v}_t| - |\mathbf{u}_t| \rangle_{\Gamma_C} \geq (\mathbf{f}, \mathbf{v}-\mathbf{u})_\Omega + (\mathbf{g}_N, \mathbf{v}-\mathbf{u})_{\Gamma_N} \quad (4.21)$$

holds for all $\mathbf{v} \in (\mathbf{u}_D + \mathbf{K}_{g_C})$.

The variational inequality in the formulation of the dual problem with Coulomb friction remains the same as for the frictionless problem or the problem with given friction. However, since the admissible set depends on the solution, the dual formulation also has to be considered in a formal sense:

Find $\boldsymbol{\sigma} \in \mathcal{K}^*_{(\boldsymbol{\sigma}\cdot\mathbf{n})_n}$ such that

$$(\mathcal{A}\boldsymbol{\sigma}, \boldsymbol{\tau}-\boldsymbol{\sigma})_\Omega \geq \langle (\boldsymbol{\tau}-\boldsymbol{\sigma})\cdot\mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle ((\boldsymbol{\tau}-\boldsymbol{\sigma})\cdot\mathbf{n})_n, g_C \rangle_{\Gamma_C} \quad (4.22)$$

holds for all $\boldsymbol{\tau} \in \mathcal{K}^*_{(\boldsymbol{\sigma}\cdot\mathbf{n})_n}$.

Since the solutions of the problem with Coulomb friction solve a specific problem with given friction, the duality result (4.1) by definition holds also for the problem with Coulomb friction:

Theorem 4.4.

(i) Let $\mathbf{u}$ *be a solution of (4.21), then* $\boldsymbol{\sigma} := \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u})$ *solves (4.22).*

(ii) Let conversely $\boldsymbol{\sigma}$ *be a solution of (4.22), then there exists* $\mathbf{u} \in H^1(\Omega)$ *with* $\mathcal{A}\boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{u})$ *such that* $\mathbf{u}$ *solves (4.21).*

4.2.1 Existence and Uniqueness

The well posedness of the problem with Coulomb friction is tightly linked to the magnitude of the friction coefficient. The well posedness of the primal problem has been studied in [Jar83, EJ98, HHNL88, Ren06]. Under the assumption of sufficiently smooth data the existence of a solution is guaranteed if the friction coefficient is small enough. More precisely the following sufficient conditions for the friction coefficient are given in [EJ98]:

$$\mu_F < \begin{cases} \sqrt{1 - \frac{1}{4}\left(\frac{\mu}{\lambda+\mu}\right)^2}, & d = 2 \\ \sqrt{\frac{\lambda+3\mu}{2\lambda+4\mu}}, & d = 3 \end{cases} \quad (4.23)$$

While there exist counter examples that show a bifurcation for large friction coefficients (see [Hil03]) the uniqueness of a solution can be shown for small friction coefficients if

it satisfies certain regularity requirements, which will be touched on in section 4.5 (see [Ren06]).

Every result on the well posedness of the primal problem with Coulomb friction can be translated into an equivalent statement on the well posedness of the dual problem with Coulomb friction due to Theorem 4.4.

4.3 Finite Element Discretization

We will now discuss the extension of the discretization in section 3.3 to the problem with given friction (4.7). Essentially we need to find an appropriate discretization of the friction condition (4.6). The linear sign condition (3.1b) for the frictionless problem allowed various weak discrete formulations (3.16) depending on the space P_h of the enforcing Lagrange multiplier. This approach, namely using a discretization $\mathbf{Q}_h$ of $\mathbf{Q}$ (conforming or unconforming), and using a weak formulation analogous to (4.6), i. e.

$$\langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_t, \mathbf{r}_h \rangle_{\Gamma_C,h} + \langle \mu_F g_F, |\mathbf{r}_h| \rangle_{\Gamma_C,h} \leq 0 \quad \forall\, \mathbf{r}_h \in \mathbf{Q}_h\,, \tag{4.24}$$

is very hard to handle both numerically and theoretically due to the non linearity in the test-functions. We will see below that for the discretized problem both $(\boldsymbol{\tau}_h \cdot \mathbf{n})_t \in \mathbf{L}^2(\Gamma_C)$ and $g_F \in L^2(\Gamma_C)$ holds. Thus the natural choice is to retreat to the strong friction condition (4.5):

$$|(\boldsymbol{\tau}_h \cdot \mathbf{n})_t| - \mu_F |g_F| \leq 0 \text{ on } \Gamma_C\,. \tag{4.25}$$

To continue in the framework of section 3.3 we write (4.25) (at least formally) in terms of a Lagrange multiplier

$$\langle |(\boldsymbol{\tau}_h \cdot \mathbf{n})_t|, r_h \rangle_{\Gamma_C,h} - \langle \mu_F |g_F|, r_h \rangle_{\Gamma_C,h} \leq 0 \quad \forall\, r_h \in Q_h\,, \tag{4.26}$$

and define the discrete admissible set

$$\begin{aligned} \boldsymbol{\mathcal{K}}^*_{g_F,h} := \{ \boldsymbol{\tau}_h \in \boldsymbol{\Sigma}^F_{\mathbf{g}_N,h} : (\operatorname{div} \boldsymbol{\tau}_h, \mathbf{v}_h) + (\mathbf{f}, \mathbf{v}_h) &= 0 \ \forall\, \mathbf{v}_h \in \mathbf{U}_h, \\ (\mathbf{as}\, \boldsymbol{\tau}_h, \boldsymbol{\gamma}_h) &= 0 \ \forall\, \boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h, \\ \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} &\leq 0 \ \forall\, w_h \in P_h, \\ \langle |(\boldsymbol{\tau}_h \cdot \mathbf{n})_t|, r_h \rangle_{\Gamma_C,h} - \langle \mu_F |g_F|, r_h \rangle_{\Gamma_C,h} &\leq 0 \ \forall\, r_h \in Q_h \}\,. \end{aligned} \tag{4.27}$$

The *discrete problem with given friction* then consists in finding $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{K}}^*_{g_F,h}$ such that

$$\mathcal{J}^C_{du}(\boldsymbol{\sigma}_h) = \min_{\boldsymbol{\tau}_h \in \boldsymbol{\mathcal{K}}^*_{g_F,h}} \mathcal{J}^C_{du}(\boldsymbol{\tau}_h)\,. \tag{4.28}$$

Since $\boldsymbol{\mathcal{K}}^*_{g_F,h}$ is again non-empty, convex and closed, Theorem 1.12 yields existence and uniqueness of the solution. Due to the nonsmoothness of the friction constraint we cannot apply standard methods for the formulation of the optimality conditions. However, we can formally still write down a KKT system for this problem, which will be justified in section 4.4:

Find $\boldsymbol{\sigma}_h \in (\boldsymbol{\Sigma}^F_{\mathbf{g}_N,h})$, $\mathbf{u}_h \in \mathbf{U}_h$, $\boldsymbol{\theta}_h \in \boldsymbol{\Theta}_h$, $p_h \in P_h$ and $q_h \in Q_h$ such that

$$\begin{aligned}
(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + (\mathbf{u}_h, \operatorname{div} \boldsymbol{\tau}_h) + \\
(\boldsymbol{\theta}_h, \mathbf{as}\, \boldsymbol{\tau}_h) + \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, p_h \rangle_{\Gamma_C,h} + \\
\langle \frac{(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t}{|(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t|} (\boldsymbol{\tau}_h \cdot \mathbf{n})_t, q_h \rangle_{\Gamma_C,h} &= \langle \boldsymbol{\tau}_h \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle (\boldsymbol{\tau}_h \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C}, \\
(\operatorname{div} \boldsymbol{\sigma}_h, \mathbf{v}_h) &= -(\mathbf{f}, \mathbf{v}_h)\ , \\
(\mathbf{as}\, \boldsymbol{\sigma}_h, \boldsymbol{\gamma}_h) &= 0\ , \\
\langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, w_h \rangle_{\Gamma_C,h} &\leq 0\ , \\
\langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, p_h \rangle_{\Gamma_C,h} &= 0\ , \\
\langle |(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t|, r_h \rangle_{\Gamma_C,h} - \langle \mu_F |g_F|, r_h \rangle_{\Gamma_C,h} &\leq 0\ , \\
\langle |(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t|, q_h \rangle_{\Gamma_C,h} - \langle \mu_F |g_F|, q_h \rangle_{\Gamma_C,h} &= 0\ ,
\end{aligned} \tag{4.29}$$

holds for all $\boldsymbol{\tau}_h \in \boldsymbol{\Sigma}^F_{\mathbf{0},h}$, $\mathbf{v}_h \in \mathbf{U}_h$, $\boldsymbol{\gamma}_h \in \boldsymbol{\Theta}_h$, $w_h \in P_h$ and $r_h \in Q_h$.

Assuming that P_h is such that the strong sign condition holds for the solution, i.e. $(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n \leq 0$, we can define a discrete counterpart of the mapping in (4.19):

$$\Psi_h : \begin{cases} \mathcal{DP}_1^-(\mathcal{S}_{h,C}) \to \mathcal{DP}_1^-(\mathcal{S}_{h,C}) \\ g_F \mapsto \Psi_h(g_F) = (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n \end{cases} \tag{4.30}$$

and in analogy to Definition 4.3 relate fixed points of this mapping to solutions of the *discrete problem with Coulomb friction*:

Find $\boldsymbol{\sigma}_h \in \boldsymbol{\mathcal{K}}^*_{(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n,h}$ such that

$$(\mathcal{A}\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h - \boldsymbol{\sigma}_h)_\Omega \geq \langle (\boldsymbol{\tau}_h - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{g}_D \rangle_\Gamma + \langle ((\boldsymbol{\tau}_h - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_n, g_C \rangle_{\Gamma_C} \tag{4.31}$$

holds for all $\boldsymbol{\tau}_h \in \boldsymbol{\mathcal{K}}^*_{(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n,h}$.

The strategy to obtain a solution to the discrete problem with Coulomb friction then consists in the following fixed point iteration:

$$[(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n]^{(k+1)} = \Psi_h([(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n]^{(k)}) \tag{4.32}$$

Consequently we will assume $g_F \in \mathcal{DP}_1^-(\mathcal{S}_{h,C}) \subset (L^2(\Gamma_C) \cap P')$ in the following.

4.3.1 Discretization of the friction condition

We could now proceed as for the sign condition in section 3.3.1 and examine different choices of Q_h, however our numerical experiments have shown that only a strong enforcement of the friction condition leads to plausible results. We have seen in section 3.3 that it is possible to enforce the strong sign condition, using a piecewise duality product (corresponding in the end to the L_2 scalar product) and piecewise polynomial Lagrange multipliers. This is no longer possible for the friction condition due to the nonlinearity

of the absolute value. However, the direct control of the values of $|(\boldsymbol{\tau}_h \cdot \mathbf{n})_t|$ at the edge-endpoints in 2D and at the face-corners in 3D allows for the strong enforcement of the friction condition (4.25), due to the linearity of $(\boldsymbol{\tau}_h \cdot \mathbf{n})_t$ on each edge/face and the convexity of the absolute value/Euclidean norm. Thus, for the implementation it will again be of great advantage to use an $\mathcal{RT}$ basis that has a property similar to (3.46), such that the related coefficients in the basis-representation of $\boldsymbol{\tau}_h$ correspond to the values of $(\boldsymbol{\tau}_h \cdot \mathbf{n})_n$ and $(\boldsymbol{\tau}_h \cdot \mathbf{n})_t$ at the edge-endpoints in 2D and at the face-corners in 3D (see section 4.4.2 for details).

The biggest disadvantage of this treatment of the friction constraint is that $\langle \cdot, \cdot \rangle_{\Gamma_C,h}$ can no longer be interpreted as an L^2-product and Q_h consists no longer of piecewise polynomials. Instead the duality product must be interpreted in a discrete pointwise sense and the Lagrange multiplier r_h is only a lumped approximation of $|\mathbf{u}_t|$. While the conformity of the stress approximation with respect to the contact conditions is preserved by this pointwise treatment of the friction condition, it severly complicates the derivation of a-priori error estimates, due to the lack of appropriate interpolation operators (cf. the discussion at the end of section 3.3.2). Furthermore the analysis of the discrete fixed point problem

$$\Psi_h(g_F) = g_F \,, \tag{4.33}$$

as is done for the primal formulation in [HHN96], is unfortunately hardly possible. However, in our numerical experiments we observe convergence of the fixed point iteration (4.32) toward a solution of (4.33) if the resolution of the contact zone is sufficiently high (see section 4.6).

4.4 Solution Method

The semismooth Newton method applied in the frictionless case can also be used to solve (4.28). However there are some important differences to consider. Since the friction condition involves the absolute value or Euclidean norm of the shear stress in the contact zone, additional nonsmooth nonlinearities need to be taken care of. We will see that the fact that the absolute value is not differentiable at the origin doesn't present a big problem in practice, despite requiring more sophisticated theory. However, the nonlinear nature of the Euclidean norm (which needs to be used in a three-dimensional setting) significantly changes the nature of the semismooth Newton method, requring more care in the implementation and slowing down its convergence.

Before we will discuss the changes in the semismooth Newton method, we need to introduce some tools from convex analysis and nonsmooth optimization. We refer to [Cla13] for a detailed treatment of this topic and present only a minimal excerpt of definitions and results that are needed for the sound treatment of (4.28).

4.4.1 Subdifferential & generalized KKT conditions

Since the constraint on the frictional shear stress involves a nonsmooth function, standard results on KKT systems cannot be applied to derive optimality conditions. How-

ever, similar results hold for convex functions. First we need a generalization of the derivative for convex functions: the subdifferential. The following can be found in [Cla13, Section 4.1].

Definition 4.5. Let $f : X \to \mathbb{R}$ be a given function, where X is a normed space, and $x \in X$. An element ζ of X^* (the dual space of X) is called a *subgradient* of f at x if it satisfies the following subgradient inequality:

$$f(y) - f(x) \geq \langle \zeta, y - x \rangle \quad \forall\, y \in X.$$

The set of all subgradients of f at x is denoted by $\partial f(x)$, and referred to as the *subdifferential* of f at x.

It is evident from this definition that x is a minimizer of f, if and only if $0 \in \partial f$. We now review two basic properties of the subdifferential.

Proposition 4.6. *Let X be a normed space, $f : X \to \mathbb{R}$ be a convex function and $x \in X$. Then*

$$\partial f(x) = \{\zeta \in X^* : f'(x; v) \geq \langle \zeta, v \rangle \ \forall\, v \in X\},$$

holds, with $f'(x; v)$ denoting the directional derivative of f at x in direction v.

Proof. See [Cla13, Proposition 4.3]. □

This proposition implies that if f is differentiable at x, the subdifferential coincides with the derivative of f at x: $\partial f(x) = \{f'(x)\}$.

Proposition 4.7. *Let X be a normed space, $f, g : X \to \mathbb{R}$ be convex functions and let there be a point $x \in X$ where f is continuous. Then we have*

$$\partial(f + g)(x) = \partial f(x) + \partial g(x) \quad \forall\, x \in X\,. \tag{4.34}$$

Proof. See [Cla13, Theorem 4.10]. □

It is evident from the definition of the subdifferential, that it is also compatible with the multiplication of a scalar. Before we arrive at the generalized optimality conditions, we will shortly illustrate the nature of the subdifferential by a simple example.

Example 4.8 (Subdifferential of the Euclidean norm)**.** Let $X = \mathbb{R}^d$ and f be the function $f(x) = |x|$. Then f is differentiable for $x \neq 0$ and we have $\partial f(x) = \{f'(x)\} = \{\frac{x}{|x|}\}$. On the other hand if $x = 0$ we have $f'(0; v) = |v|$. Observing

$$\langle \zeta, v \rangle = v^{\mathsf{T}} \zeta \leq |v||\zeta| \leq |v| = f'(0, v) \Leftrightarrow |\zeta| \leq 1$$

and applying Proposition 4.6 leads to

$$\partial f(x) = \partial |x| = \begin{cases} \{\frac{x}{|x|}\} & \text{if } x \neq 0 \\ \overline{B}_1^d(0) & \text{if } x = 0 \end{cases}, \tag{4.35}$$

where $\overline{B}_r^d(y)$ is again the d-dimensional closed ball around y with radius r.

We can now recapitulate the main result on KKT systems of convex (but not necessarily smooth) optimization problems:

Theorem 4.9. *Let X be a normed space and $f : X \to \mathbb{R}$ be a continuous and convex function. Let $g : X \to \mathbb{R}^m$ and $h : X \to \mathbb{R}^n$ be continuous functions such that g_i $(i = 1, 2, ..., m)$ are convex and h_j $(j = 1, 2, ..., n)$ are affine. Let x_* be a solution of the following problem:*

$$\text{Minimize } f(x) \text{ subject to } g(x) \leq 0,\ h(x) = 0,\ x \in X, \tag{4.36}$$

then there exist Lagrange multipliers $(\eta, \gamma, \lambda) \in \mathbb{R}^m \times \mathbb{R}^n$ satisfying the following conditions:

$$0 \in \partial(\eta f(x_*) + \langle \gamma, g(x_*)\rangle + \langle \lambda, h(x_*)\rangle)\,, \tag{4.37a}$$
$$(\eta, \gamma, \lambda) \neq 0\,, \tag{4.37b}$$
$$\eta \in \{0, 1\}\,, \tag{4.37c}$$
$$\gamma \geq 0\,, \tag{4.37d}$$
$$\langle \gamma, g(x_*)\rangle = 0\,. \tag{4.37e}$$

Proof. The proof is essentially that of [Cla13, Theorem 9.4]. The only thing left to show is that, for f, g, h continuous and convex on X, (4.37a) is equivalent to

$$\eta f(x) + \langle \gamma, g(x)\rangle + \langle \lambda, h(x)\rangle \geq \eta f(x_*) \quad \forall\, x \in X\,.$$

Indeed, since $h(x_*) = 0$ and (4.37e) hold, we have

$$\eta f(x_*) = \eta f(x_*) + \langle \gamma, g(x_*)\rangle + \langle \lambda, h(x_*)\rangle\,.$$

Consequently setting $\hat{f}(x) := (\eta f(x) + \langle \gamma, g(x)\rangle + \langle \lambda, h(x)\rangle)$ and recalling Definition 4.5 we obtain:

$$0 \in \partial \hat{f}(x_*) \Leftrightarrow \hat{f}(x) \geq \hat{f}(x_*) \quad \forall\, x \in X\,. \tag{4.38}$$

□

Similarly to the smooth case, if the constraints satisfy additional requirements, these generalized KKT conditions are also sufficient for optimality:

Theorem 4.10 (Slater condition)**.** *Let x_* be admissible for (4.36) and suppose there exist associated Lagrange multipliers (η, γ, λ) that satisfy (4.37). Let the affine functions of the equality constraints be independent (meaning that the set $\{h'_j : j = 1, 2, ..., n\}$ is independent) and let there be a point $x^\circ \in X$ that satisfies $g(x^\circ) < 0$ and $h(x^\circ) = 0$. Then $\eta = 1$ and x_* solves (4.36).*

Proof. Again the main part of the proof can be found in [Cla13, Theorem 9.8]. What remains to be shown, is that $\eta = 1$ implies the optimality of x_*. Indeed, if $\eta = 1$, for any admissible $x \in X$ (i.e. $g(x) \leq 0$ and $h(x) = 0$) we have

$$f(x_*) = \hat{f}(x_*) \leq \hat{f}(x) = f(x) + \langle \gamma, g(x)\rangle \leq f(x)\,, \tag{4.39}$$

where we have used (4.38) and (4.37d) for the inequalities . □

4.4.2 Application to the discrete frictional contact problem

As done for the frictionless case, we can reduce all quantities in (4.27) to their representation in Euclidean space. The matrix vector notation described in (3.41) is again used with the adaptation that n_Σ is changed since the frictionless condition is no more incorporated in the ansatz space. We can thus use the concept of subdifferentials (applying also Proposition 4.7) to write down a KKT system for (4.28).

Find $x \in \mathbb{R}^{n_\Sigma}$, $y \in \mathbb{R}^{n_U}$, $z \in \mathbb{R}^{n_\Theta}$, $p \in \mathbb{R}^m$ and $q \in \mathbb{R}^m$, such that

$$\begin{aligned}
Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p + (\partial G(x))^\mathsf{T} q - b &\ni 0\,, && \text{(4.40a)}\\
Bx + f &= 0\,, && \text{(4.40b)}\\
Cx &= 0\,, && \text{(4.40c)}\\
Dx &\le 0\,, && \text{(4.40d)}\\
G(x) &\le 0\,, && \text{(4.40e)}\\
p &\ge 0\,, && \text{(4.40f)}\\
q &\ge 0\,, && \text{(4.40g)}\\
p^\mathsf{T} Dx &= 0\,, && \text{(4.40h)}\\
q^\mathsf{T} G(x) &= 0\,. && \text{(4.40i)}
\end{aligned}$$

Here $g_i(x) := |G_i x| - s_i$ is the ith component of $G(x) : \mathbb{R}^{n_\Sigma} \to \mathbb{R}^m$, with $s_i = \mu_F |g_F|_{S_{l_i}}(\mathbf{z}_{k_i})| \ge 0$ and $G_i \in \mathbb{R}^{(d-1)\times n_\Sigma}$ being a linear mapping such that $G_i x = (\boldsymbol{\sigma}_h|_{S_{l_i}}(\mathbf{z}_{k_i}) \cdot \mathbf{n})_t$. The indices l_i and k_i represent a suitable numbering of the sides S and nodes $\mathbf{z}$ on Γ_C corresponding to a relationship analogous to the one in (3.46). Accordingly

$$\partial g_i(x) = \partial |G_i x| = \begin{cases} \{\nabla |G_i x|\} = \left\{ \frac{x^\mathsf{T} G_i^\mathsf{T} G_i}{|G_i x|} \right\} & \text{if } G_i x \neq 0 \\ \{v^\mathsf{T} G_i : v \in \overline{B}_1^{(d-1)}(0) & \text{if } G_i x = 0 \end{cases} \tag{4.41}$$

is the ith line of the subdifferential $\partial G(x)$.

Analogous to (3.51) the relations (4.40d) to (4.40i) can be summarized into nonsmooth equations and we obtain the following nonsmooth system:

Find $x \in \mathbb{R}^{n_\Sigma}$, $y \in \mathbb{R}^{n_U}$, $z \in \mathbb{R}^{n_\Theta}$, $p \in \mathbb{R}^m$ and $q \in \mathbb{R}^m$, such that

$$\begin{aligned}
Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p + (\partial G(x))^\mathsf{T} q - b &\ni 0\,, && \text{(4.42a)}\\
Bx + f &= 0\,, && \text{(4.42b)}\\
Cx &= 0\,, && \text{(4.42c)}\\
p - \max\{0, p + \kappa(Dx)\} &= 0\,, && \text{(4.42d)}\\
q - \max\{0, q + \kappa G(x)\} &= 0\,, && \text{(4.42e)}
\end{aligned}$$

hold.

We now want to examine whether (4.40) really qualifies for an application of Theorems 4.9 and 4.10. The energy functional $\mathcal{J}_{du}^C$ is convex and continuous and the constraint

functions also satisfy the assumptions of Theorem 4.9. However the Slater condition does not hold if there is an i with $s_i = 0$. Before we take care of that, we want to point out that really the major difference to (3.51) is the appearance of the subdifferential in (4.42a) requiring the first line to be written as a set relation rather than an equation. However, a closer look reveals that in practice we can avoid having to deal with subdifferentials altogether.

For this purpose let $I_0 := \{i = 1, 2, ..., m : s_i = 0\}$ be the set of indices where the given contact pressure vanishes. For these indices the friction condition reduces to a homogeneous boundary condition for the shear stress (frictionless condition) which can again be incorporated into the ansatz space $\tilde{\boldsymbol{\Sigma}}_h := \boldsymbol{\Sigma}_h \setminus \{\boldsymbol{\psi}_h : (\boldsymbol{\psi}_h|_{S_{l_i}}(\mathbf{z}_{k_i})_T \cdot \mathbf{n}) \neq 0, i \in I_0\}$ by eliminating the corresponding basis functions. The remaining friction constraints can be expressed by $\tilde{G}(\tilde{x}) : \mathbb{R}^{\tilde{n}_\Sigma} \to \mathbb{R}^{\tilde{m}}$ with $\tilde{n}_\Sigma := \dim(\tilde{\boldsymbol{\Sigma}}_h)$ and $\tilde{m} := |I_>| := |\{i = 1, 2, ..., m : s_i > 0\}| = m - |I_0|$.

First of all, this modification guarantees that the Slater condition holds and thus the solution of (4.42) exists and solves (4.28). Indeed the inf-sup condition (2.21) yields the existence of a $\boldsymbol{\tau}_h^\circ \in \boldsymbol{\Sigma}_{\mathbf{g}_N,h}^F$ with $(\boldsymbol{\tau}_h^\circ \cdot \mathbf{n})_n = -1$ and $(\boldsymbol{\tau}_h^\circ \cdot \mathbf{n})_t = 0$ on Γ_C such that (4.40b) and (4.40c) hold. This $\boldsymbol{\tau}_h^\circ$ corresponds to a x° that satisfies the assumptions of Theorem 4.10 for the modified friction constraint. (Since the case $g_F \equiv 0$ corresponds to the frictionless case, this argument also yields the well-posedness of (3.51).)

Furthermore, if we formally extend each gradient $\nabla|\tilde{G}_i\tilde{x}|$ to the origin by some arbitrary vector $\nabla|0| := v^\mathsf{T} \in \mathbb{R}^{\tilde{n}_\Sigma}$, the complementarity condition $\tilde{q}^\mathsf{T}\tilde{G}(\tilde{x}) = 0$ together with $s_i > 0$ for all $i \in I_>$ implies $\tilde{q}^\mathsf{T}\partial\tilde{G}(\tilde{x}) = \tilde{q}^\mathsf{T}\nabla\tilde{G}(\tilde{x})$ since $\tilde{q}_i = 0$ whenever $\partial|\tilde{G}_i\tilde{x}| \neq \nabla|\tilde{G}_i\tilde{x}|$.

To simplify notation below we will omit the $\tilde{}$ on the variables x and q as well as on the changed matrices and keep it only on $\tilde{G}$, $\tilde{n}_\Sigma$ and $\tilde{m}$ to remind the reader that the constraints have been reduced. With this in mind we can avoid dealing with sets and rewrite (4.42) in the following way:

Find $x \in \mathbb{R}^{\tilde{n}_\Sigma}$, $y \in \mathbb{R}^{n_U}$, $z \in \mathbb{R}^{n_\Theta}$, $p \in \mathbb{R}^m$ and $q \in \mathbb{R}^{\tilde{m}}$, such that

$$Ax + B^\mathsf{T}y + C^\mathsf{T}z + D^\mathsf{T}p + (\nabla\tilde{G}(x))^\mathsf{T}q - b = 0\,, \tag{4.43a}$$

$$Bx - f = 0\,, \tag{4.43b}$$

$$Cx = 0\,, \tag{4.43c}$$

$$p - \max\{0, p + \kappa(Dx)\} = 0\,, \tag{4.43d}$$

$$q - \max\{0, q + \kappa\tilde{G}(x)\} = 0 \tag{4.43e}$$

hold.

4.4.3 Semismooth Newton method

Even though we were able to express (4.28) as a system of nonlinear equations there seems to remain another obstacle to the application of the semismooth Newton method. As shown in chapter 3 the maximum function and the Euclidean norm are both Newton-differentiable as well as their composition. Thus equations (4.43d) and (4.43e) are not the problem. The term in $(\nabla\tilde{G}(x))^\mathsf{T}q$ in (4.43a), however, is not Newton differentiable at $x = 0$.

The first idea to tackle this problem would be to apply some kind of regularization to the Euclidean norm (such as used for example in [KO88, Chapter 10]) which takes care of the singularity at the origin. We shall see, however, that if the regularization is restricted to a small enough ball around the origin, it has no effect on the iterates of the resultant semismooth Newton method, and consequently can be omitted. To more clearly illustrate this point, we consider a simplified model problem which nonetheless exhibits the essential features of (4.43).

Example 4.11. Let $A \in \mathbb{R}^{2\times 2}$ s.p.d. and $b \in \mathbb{R}^2$ be given and let $m = 1$, $s_1 = 1$ and $G_1 = I$ (the identity matrix in $\mathbb{R}^{2\times 2}$) determine $G(x) = g_1(x)$ as above. Our model problem is then to find $x \in \mathbb{R}^2$ and $q \in \mathbb{R}$, that satisfy

$$F(x,q) := \begin{pmatrix} Ax + (\nabla G(x))^{\mathsf{T}} q - b \\ q - \max\{0, q + \kappa G(x)\} \end{pmatrix} = \begin{pmatrix} Ax + \frac{x}{|x|} q - b \\ q - \max\{0, q + \kappa(|x| - 1)\} \end{pmatrix} = 0. \tag{4.44}$$

For $x \neq 0$

$$DF(x,q) = \begin{pmatrix} A + \nabla^2 G(x) q & \frac{x}{|x|} \\ \kappa_S \frac{x^{\mathsf{T}}}{|x|} & 1_{S^c} \end{pmatrix}, \tag{4.45}$$

is a Newton derivative of $F(x,q)$ whenever $x \neq 0$, with

$$S := \{i \in \{1, ..., m\} : q_i + \kappa g_i(x) > 0\} \tag{4.46}$$

being the active set, S^c its complement and the subscript indicating the vanishing of this term dependent on S as was elaborated in (3.63). In our simple example either $S = \{1\}$ or $S = \emptyset$ holds. The Hessian of G in (4.45) reads:

$$\nabla^2 G(x) = \frac{1}{|x|^3} \begin{pmatrix} x_2^2 & x_1 x_2 \\ x_1 x_2 & x_1^2 \end{pmatrix}. \tag{4.47}$$

The Newton step $\delta_k := \begin{pmatrix} \delta x_k \\ \delta q_k \end{pmatrix}$ then solves one of the following systems dependent on S:

$$\begin{pmatrix} A + \nabla^2 G(x_k) q_k & \frac{x_k}{|x_k|} \\ \kappa \frac{x_k^{\mathsf{T}}}{|x_k|} & 0 \end{pmatrix} \begin{pmatrix} \delta x_k \\ \delta q_k \end{pmatrix} = - \begin{pmatrix} Ax_k + \frac{x_k}{|x_k|} q_k - b \\ \kappa(|x_k| - 1) \end{pmatrix} \qquad \text{if } S_k \neq \emptyset \tag{4.48}$$

$$\begin{pmatrix} A + \nabla^2 G(x_k) q_k & \frac{x_k}{|x_k|} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \delta x_k \\ \delta q_k \end{pmatrix} = - \begin{pmatrix} Ax_k + \frac{x_k}{|x_k|} q_k - b \\ q_k \end{pmatrix} \qquad \text{if } S_k = \emptyset \tag{4.49}$$

The second line in (4.48) guarantees $|x_{k+1}| := |x_k + \delta x_k| \geq 1$, since, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |x_k + \delta x_k|^2 &= |x_k|^2 + |\delta x_k|^2 + 2x_k^{\mathsf{T}} \delta x_k = |x_k|^2 + \left| \frac{x_k}{|x_k|} \right|^2 |\delta x_k|^2 + 2x_k^{\mathsf{T}} \delta x_k \\ &\geq |x_k|^2 + \left(\frac{x_k^{\mathsf{T}} \delta x_k}{|x_k|} \right)^2 + 2x_k^{\mathsf{T}} \delta x_k \\ &= |x_k|^2 + (|x_k| - 1)^2 - 2|x_k|(|x_k| - 1) = 1 \,. \end{aligned}$$

On the other hand the second line in (4.49) guarantees $q_{k+1} := q_k + \delta q_k = 0$. Together this implies that if the initial value q_0 is set to zero we have

$$|x_k| < 1 \Rightarrow S_{k-1} = \emptyset \Rightarrow q_k = 0 \quad \text{for } k \geq 1 \,, \tag{4.50}$$

thus making the algorithm blind for any regularization of the constraint close enough to the origin.

The same reasoning applies when using the semismooth Newton method to solve (4.43):

Theorem 4.12. *Let*

$$|x|_\varepsilon := \begin{cases} |x| & \text{if } |x| > \varepsilon \\ \frac{x^\mathsf{T} x}{2\varepsilon} + \frac{\varepsilon}{2} & \text{if } |x| \leq \varepsilon \end{cases} \tag{4.51}$$

define a regularizaton of the Euclidean norm for each $\varepsilon > 0$. If $q_0 = 0$ then the sequences $\{x_k\}$,$\{y_k\}$,$\{z_k\}$,$\{p_k\}$ and $\{q_k\}$ produced by the semismooth Newton method applied to

$$Ax + B^\mathsf{T} y + C^\mathsf{T} z + D^\mathsf{T} p + (\nabla \tilde{G}_\varepsilon(x))^\mathsf{T} q - b = 0 \,, \tag{4.52a}$$

$$Bx - f = 0 \,, \tag{4.52b}$$

$$Cx = 0 \,, \tag{4.52c}$$

$$p - \max\{0, p + \kappa(Dx)\} = 0 \,, \tag{4.52d}$$

$$q - \max\{0, q + \kappa \tilde{G}_\varepsilon(x)\} = 0 \tag{4.52e}$$

are the same for all $\varepsilon < \min\{s_i : i \in I_>\}$, with $g_i^\varepsilon(x) := |G_i x|_\varepsilon - s_i$ being the components of $\tilde{G}_\varepsilon(x)$. Furthermore, if $\varepsilon < \min\{s_i : i \in I_>\}$, then (4.52) and (4.43) are equivalent.

Proof. Let $\varepsilon < \min\{s_i : i \in I_>\}$ hold, then $g_i^\varepsilon(x) = g_i(x)$ holds, whenever $|G_i x| \geq s_i$.

As in example 4.11 the structure of the Newton problem together with $q_0 = 0$ implies

$$|G_i x_k| < s_i \Rightarrow [q_k]_i = 0 \quad \text{for } k \geq 0, \text{ for } i \in I_> \,, \tag{4.53}$$

thus we have

$$g_i^\varepsilon(x_k) \neq g_i(x_k) \Rightarrow [q_k]_i = 0 \quad \text{for } k \geq 0, \text{ for } i \in I_> \,. \tag{4.54}$$

Since all occurences in the Newton system of $g_i^\varepsilon(x_k)$ and its derivatives are multiplied with $[q_k]_i$, (4.54) implies the first assertion.

The second assertion is a consequence of (4.52e) and (4.43e) being equivalent. Indeed, $g_i^\varepsilon(x)$ and $g_i(x)$ always have the same sign, thus the same arguments as in Lemma 3.8 apply. Consequently $(\nabla \tilde{G}_\varepsilon(x))^\mathsf{T} q = (\nabla \tilde{G}(x))^\mathsf{T} q$, since $q_i = 0$ whenever $\nabla g_i^\varepsilon(x) \neq \nabla g_i(x)$. □

Remark. The assumption regarding the initial value, $q_0 = 0$, can be weakend by requiring only those components of q_0 to vanish, where the corresponding constraint does not satisfy $g_i^\varepsilon(x_0) = g_i(x_0)$, or, more practically, by enforcing the initial values to satisfy

$$g_i(x_0) < 0 \Rightarrow [q_0]_i = 0 \,. \tag{4.55}$$

Being careful about the inital values thus allows us to apply the semismooth Newton method to (4.43) despite the non-Newton-differentiable term $(\nabla \tilde{G}(x))^\mathsf{T} q$.

In our numerical experiments we have found that, while (3.59) performs very well for the frictionless case in 2D and 3D, as well as for the frictional case in 2D, taking the undamped Newton step in the frictional case in 3D often leads to the algorithm getting caught in a loop. As Example 4.11 illustrates, the friction constraint in the three dimensional case is of highly nonlinear nature, while in 2D it is piecewise linear. This is the reason why in 3D it is not enough anymore to find the optimal active set S (otherwise (4.44) would maximally take 4 steps to be solved), and probably also the reason why the algorithm doesn't converge without further adjustments. Remembering that Theorem 3.13 only provides only a local convergence result, we apply an inexact line search algorithm (damping) to obtain global convergence.

4.4.4 Damping

The idea behind damping in a Newton setting is the interpretation of the Newton step as a descent direction. Without going into too much detail, any solution of the equation $F(\xi) = 0$ is also a solution of the minimization problem

$$\min_{\xi} f(\xi) \tag{4.56}$$

with $f(\xi) := \|F(\xi)\|^2$. If F is differentiable at ξ_k and $\det(F'(\xi_k)) \neq 0$, the definition of the Newton step (3.59a) guarantees the existence of an $\alpha > 0$ such that $f(\xi_{k+1}) < f(\xi_k)$ with $\xi_{k+1} := \xi_k + \alpha\delta\xi_k$. Indeed for α small enough, we have

$$\begin{aligned} f(\xi_{k+1}) &= f(\xi_k) + \alpha\nabla f(\xi_k)^{\mathsf{T}}\delta\xi_k + o(\alpha\delta\xi_k) \\ &= f(\xi_k) + 2\alpha F(\xi_k)^{\mathsf{T}}F'(\xi_k)\delta\xi_k + o(\alpha\delta\xi_k) \\ &= f(\xi_k) - 2\alpha\|F(\xi_k)\|^2 + o(\alpha\delta\xi_k) < f(\xi_k)\,. \end{aligned}$$

This gives rise to a simple damped Newton algorithm:

$$\begin{aligned} &\text{Reduce } \alpha \text{ until } f(x_k + \alpha\delta x_k) < f(x_k)\,, \\ &\text{Set } x_{k+1} := x_k + \alpha\delta x_k\,. \end{aligned} \tag{4.57}$$

One could try to further improve the convergence behaviour by using more sophisticated line search approaches such as the Armijo or Wolfe conditions (see [NW06, Chapter 3]).

Remark. Despite the theoretical existence of $\alpha > 0$, in numerical computations sometimes no measurable decrease of f is achieved due to the finite precision of floating point arithmetics. When this occurs, taking the full Newton step to "restart" the algorithm often helps to achieve convergence.

Once we apply such a scheme (that requires a certain decrease of the functional $\|F(\xi)\|^2$ at every iteration) to our problem (4.43), the parameter κ in (4.43d) and (4.43e) starts to play a bigger role. Besides balancing the multipliers and the constraints within (4.43d) and (4.43e) (as remarked after Lemma (3.8)) it also acts as a weighting of the constraint violation connected to (4.43d) and (4.43e) within the total residual $\|F(\xi)\|^2 := \|F(x, y, z, p, q)\|^2$.

In other words, if κ is chosen large, every step will be required to significantly reduce the feasability error expressed by $Dx_k > 0$ and $\tilde{G}(x_k) > 0$ while the reduction of the other residuals (4.43a)-(4.43c) will be less relevant to the decision whether or not a step will be accepted. If κ is chosen small the opposite behaviour is favoured. Even in a simple problem, such as Example 4.11, it can be observed that both extremes (κ too large or too small) can severely slow down or even prevent convergence. Thus the correct choice of κ becomes an important issue.

We will conclude this section with a few thoughts on how to choose κ. Analogously to refering to the complete system (4.43) as

$$F(\xi) = F(x, y, z, p, q) = 0$$

we will refer to equations (4.43a) to (4.43e) as $F_1(x, y, z, p, q)$ to $F_5(x, y, z, p, q)$. Since F_2 and F_3 are linear, their residual will vanish for all $k > k^*$ if it has vanished for one k^*. Our focus thus rests on balancing the resdiuals of (4.43a), (4.43d) and (4.43e). Since for $|\delta x_k| \ll |x_k|$ we have $\|\nabla\tilde{G}(x_k + \delta x_k) - \nabla\tilde{G}(x_k)\|_2 \approx \frac{|\delta x_k|}{|x_k|}$, we also have

$$\|F_1(x_k + \delta x_k, y_k, z_k, p_k, q_k) - F_1(\xi_k)\|_2 \approx \|A\delta x_k\|_2 \ .$$

Since A is s.p.d. the change in the residual of F_1 can be estimated by the eigenvalues of A:

$$\lambda_{\min}(A)\|\delta x_k\|_2 \leq \|A\delta x_k\|_2 \leq \lambda_{\max}(A)\|\delta x_k\|_2 \tag{4.58}$$

If $[q_k]_i + \kappa g_i(x_k) > 0$ and $[q_k]_i + \kappa g_i(x_k + \delta x_k) > 0$, we have

$$\begin{aligned} |F_{5,i}(x_k + \delta x_k, y_k, z_k, p_k, q_k) - F_{5,i}(\xi_k)| &= \kappa|g_i(x_k) - g_i(x_k + \delta x_k)| \\ &= \kappa||G_i x_k| - |G_i(x_k + \delta x_k)|| \\ &\leq \kappa|G_i\delta x_k| \leq \kappa\|\delta x_k\|_2 \ , \end{aligned}$$

and for F_4 we get a similar estimate. This suggests the choice of $\kappa \in [\lambda_{\min}(A), \lambda_{\max}(A)]$ in order to achieve a balanced weighting of the residuals of F_1, F_4 and F_5. Since only the magnitude of the eigenvalues is of interest, it is enough to estimate the largest and smallest eigenvalues on the coarsest refinement level and then adjust κ as h decreases.

Even with these adaptations the convergence of the semismooth Newton method is significantely slower in the threedimensional case, taking 150-300 steps to converge, while in 2D the algorithm typically terminates after 15-30 steps.

4.5 A-posteriori error analysis

In this section we will discuss the modifications necessary to generalize the result in Theorem 3.7 to the contact problems with given friction and Coulomb friction. Since in the frictional case the shear stress does not vanish on all of Γ_C, we have to revise the treatment of the contact boundary term. While this is straightforward for the problem with given friction, a similar result for Coulomb friction can only be obtained under

additional regularity assumptions on the solution. This is related to the fact that in general the solution to the problem with Coulomb friction is not unique. We will first review the modifications necessary for the problem with given friction.

Theorem 4.13. *Let $\boldsymbol{\sigma}$ solve (4.7) and $\boldsymbol{\sigma}_h$ solve (4.28). Let $\boldsymbol{\sigma}_h$ satisfy (4.25) and (3.1b) and let $\mathbf{u}_h^R \in (\mathbf{u}_D + \mathbf{K}_{g_C})$. Then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathrm{div})} \leq C\,(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \tilde{\eta}_5) \tag{4.59}$$

holds with $\eta_1, \eta_2, \eta_3, \eta_4$ as in Theorem 3.7 and

$$\tilde{\eta}_5 := \left(\langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} - \langle \mu_F g_F, |(\mathbf{u}_h^R)_t| \rangle_{\Gamma_C} \right)^{1/2} ,$$

and a constant C which is independent of λ and h.

Proof. We examine the boundary term in (2.33) and start as in the proof of Theorem 3.7 using $\boldsymbol{\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_{\mathbf{0}}^F$ and $\mathbf{u} - \mathbf{u}_h^R \in \mathbf{V}$ to restrict it to Γ_C and separate normal and tangential components. We thus obtain

$$\begin{aligned} \langle (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{u} - \mathbf{u}_h^R \rangle_{\Gamma} &= \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_n, (\mathbf{u} - \mathbf{u}_h^R)_n \rangle_{\Gamma_C} \\ &\quad + \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_t, (\mathbf{u} - \mathbf{u}_h^R)_t \rangle_{\Gamma_C} \\ &\leq \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C \rangle_{\Gamma_C} \\ &\quad + \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_t, (\mathbf{u} - \mathbf{u}_h^R)_t \rangle_{\Gamma_C} \end{aligned} \tag{4.60}$$

where the normal component was estimated as in the proof of Theorem 3.7. We are going to estimate the tangential part by computable quantities. First we expand it:

$$\begin{aligned} \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_t, (\mathbf{u} - \mathbf{u}_h^R)_t \rangle_{\Gamma_C} &= \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} \\ &\quad + \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} \end{aligned}$$

and then, using (4.25) and (4.13), estimate the first part by

$$\begin{aligned} \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} &\leq \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} + \langle |(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t|, |\mathbf{u}_t| \rangle_{\Gamma_C} \\ &\leq \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} - \langle \mu_F g_F, |\mathbf{u}_t| \rangle_{\Gamma_C} \\ &= 0\,. \end{aligned}$$

For the second part we observe $-(\mathbf{u}_h^R)_t \in \mathbf{Q}$ and use (4.6) to obtain

$$\begin{aligned} \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} &\leq \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} \\ &\quad - \langle \mu_F g_F, |(\mathbf{u}_h^R)_t| \rangle_{\Gamma_C} . \end{aligned}$$

Thus we have

$$\begin{aligned} \langle (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \mathbf{u} - \mathbf{u}_h^R \rangle_{\Gamma} &\leq \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, (\mathbf{u}_h^R)_n - g_C \rangle_{\Gamma_C} \\ &\quad + \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} - \langle \mu_F g_F, |(\mathbf{u}_h^R)_t| \rangle_{\Gamma_C} . \end{aligned}$$

and the result follows by the arguments used in the proof of Theorem 2.4. □

If we construct $\mathbf{u}_h^R$ in a way that it satisfies the discrete frictional complementarity condition, i.e. enforcing $(\mathbf{u}_h^R)_t = 0$ on S whenever $|(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t| \neq \mu_F |g_F|$ on S for all $S \in \mathcal{S}_{h,C}$, then $\eta_5 = 0$ holds and we again obtain the same error estimator as for the problem of linear elasticity.

In order to extend this reliabilty result to the problem with Coulomb friction with sufficiently small friction parameter μ_F an inverse triangle-type inequality

$$\mu_F \| |\mathbf{u}_t| - |(\mathbf{u}_h^R)_t| \|_{1/2,\Gamma_C} \leq \gamma(\mu_F) \|\mathbf{u}_t - (\mathbf{u}_h^R)_t\|_{1/2,\Gamma_C} \tag{4.61}$$

is required to hold with γ being a continuous function which is independent of h and satisfies $\gamma(0) = 0$. Such an inequality trivially holds (with $\gamma(\mu_F) = \mu_F$ if $H^{1/2}(\Gamma_C)$ is replaced by $L^2(\Gamma_C)$ which occurs in the context of the regularized friction law in [Cap14, Chapter 8]. It can, however, not be expected to hold in $H^{1/2}(\Gamma_C)$ without making additional assumptions. Before we discuss in detail the assumptions that guarantee (4.61) we set forth its consequence in the next theorem.

Theorem 4.14. *Let $\boldsymbol{\sigma}$ solve (4.22) and $\boldsymbol{\sigma}_h$ solve (4.31). Let $\boldsymbol{\sigma}_h$ satisfy (3.1b) and*

$$|(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t| - \mu_F |(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n| \leq 0 \text{ on } \Gamma_C \,. \tag{4.62}$$

Let $\mathbf{u}_h^R \in (\mathbf{u}_D + \mathbf{K}_{g_C})$ and (4.61) hold. Then, for sufficiently small friction coefficient μ_F,

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{H(\mathrm{div})} \leq C \left(\eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5\right) \tag{4.63}$$

holds with $\eta_1, \eta_2, \eta_3, \eta_4$ as in Theorem 3.7 and

$$\eta_5 := \left(\langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} - \langle \mu_F (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, |(\mathbf{u}_h^R)_t| \rangle_{\Gamma_C} \right)^{1/2} ,$$

and a constant C which is independent of λ and h.

Proof. The start of the proof is the same as for Theorem 4.13. We therefore jump directly to the treatment of the tangential boundary term:

$$\begin{aligned} \langle ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n})_t, (\mathbf{u} - \mathbf{u}_h^R)_t \rangle_{\Gamma_C} &= \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} \\ &\quad - \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} \\ &\quad + \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} \\ &\quad - \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, (\mathbf{u}_h^R)_t \rangle_{\Gamma_C} \end{aligned} \tag{4.64}$$

but this time, using (4.62), obtain for the first part

$$\begin{aligned} \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} - \langle (\boldsymbol{\sigma}_h \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} &\leq \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} \\ &\quad + \langle |(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t|, |\mathbf{u}_t| \rangle_{\Gamma_C} \\ &\leq \langle (\boldsymbol{\sigma} \cdot \mathbf{n})_t, \mathbf{u}_t \rangle_{\Gamma_C} \\ &\quad - \langle \mu_F (\boldsymbol{\sigma}_h \cdot \mathbf{n})_n, |\mathbf{u}_t| \rangle_{\Gamma_C} . \end{aligned} \tag{4.65}$$

For the second part we again observe $-(\mathbf{u}_h^R)_t \in \mathbf{Q}$ to obtain

$$\begin{aligned}\langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_t,(\mathbf{u}_h^R)_t\rangle_{\Gamma_C}-\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,(\mathbf{u}_h^R)_t\rangle_{\Gamma_C}&\leq\langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_t,(\mathbf{u}_h^R)_t\rangle_{\Gamma_C}\\&\quad-\langle\mu_F(\boldsymbol{\sigma}\cdot\mathbf{n})_n,|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}.\end{aligned}\tag{4.66}$$

Combining (4.64) - (4.66) and adding and subtracting the right terms yields

$$\begin{aligned}\langle((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_t,(\mathbf{u}-\mathbf{u}_h^R)_t\rangle_{\Gamma_C}&\leq\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{u}_t\rangle_{\Gamma_C}\\&\quad-\langle\mu_F(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n,|\mathbf{u}_t|\rangle_{\Gamma_C}\\&\quad+\langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_t,(\mathbf{u}_h^R)_t\rangle_{\Gamma_C}\\&\quad-\langle\mu_F(\boldsymbol{\sigma}\cdot\mathbf{n})_n,|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}\\&=\langle(\boldsymbol{\sigma}\cdot\mathbf{n})_t,\mathbf{u}_t\rangle_{\Gamma_C}\\&\quad-\langle\mu_F(\boldsymbol{\sigma}\cdot\mathbf{n})_n,|\mathbf{u}_t|\rangle_{\Gamma_C}\\&\quad+\langle(\boldsymbol{\sigma}_h\cdot\mathbf{n})_t,(\mathbf{u}_h^R)_t\rangle_{\Gamma_C}\\&\quad-\langle\mu_F(\boldsymbol{\sigma}_h\cdot\mathbf{n})_n,|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}\\&\quad+\langle\mu_F((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n,|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}\\&=\eta_5^2+\langle\mu_F((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n,|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}.\end{aligned}\tag{4.67}$$

where we used the frictional complementarity condition for the last equality. Using (4.61) the last term in (4.67) can then be estimated by

$$\begin{aligned}\langle\mu_F((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n,|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|\rangle_{\Gamma_C}&\leq\mu_F\|((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n\|_{-1/2,\Gamma_C}\||\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|\|_{1/2,\Gamma_C}\\&\leq\gamma(\mu_F)\|((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n\|_{-1/2,\Gamma_C}\|\mathbf{u}_t-(\mathbf{u}_h^R)_t\|_{1/2,\Gamma_C}\\&\leq\gamma(\mu_F)\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\cdot\mathbf{n}\|_{-1/2,\Gamma}\|\mathbf{u}-\mathbf{u}_h^R\|_{1/2,\Gamma}.\end{aligned}$$

Recalling the trace inequalities (1.8) and (1.14), together with Korn's (1.54) and Young's (1.57) inequalities (with $\delta>0$ to be chosen appropriately) we obtain

$$\begin{aligned}\langle\mu_F((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n})_n,|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|,\rangle_{\Gamma_C}&\leq\frac{C_1\delta}{2}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathrm{div},\Omega}\\&\quad+\frac{C_2\gamma(\mu_F)^2}{2\delta}\|\boldsymbol{\varepsilon}(\mathbf{u}-\mathbf{u}_h^R)\|^2_{0,\Omega}.\end{aligned}\tag{4.68}$$

Combining (4.68) with (4.67) and (4.60) we obtain the following estimate for the boundary term in (2.33):

$$\langle(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n},\mathbf{u}-\mathbf{u}_h^R\rangle_\Gamma\leq\eta_4^2+\eta_5^2+\frac{C_1\delta}{2}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|^2_{\mathrm{div},\Omega}+\frac{C_2\gamma(\mu_F)^2}{2\delta}\|\boldsymbol{\varepsilon}(\mathbf{u}-\mathbf{u}_h^R)\|^2_{0,\Omega}.$$

Inserting this into (2.33) and proceeding as in the proof of Theorem 2.4 yields the result. □

4.5.1 A norm equivalence in $H^{1/2}(\Gamma)$

The purpose of this section is to show that (4.61) is indeed fulfilled under certain conditions on $(\mathbf{u}_h^R)_t$ which can be verified from the numerical experiments and under certain assumptions on $\mathbf{u}_t$ for the exact solution. We restrict ourselves to the two-dimensional case and assume that Γ_C is a connected smooth boundary segment.

Theorem 4.15. *Assume that Γ_C contains a segment Γ_C° where $(\mathbf{u}_h^R)_t \equiv 0$ and that the tangential derivative satisfies*

$$m \leq \partial_{\mathbf{t}}(\mathbf{u}_h^R)_t \leq M \text{ (or } m \leq -\partial_{\mathbf{t}}(\mathbf{u}_h^R)_t \leq M\text{) uniformly on } \Gamma_C \backslash \Gamma_C^\circ\,, \tag{4.69}$$

with positive constants m and M, and $|\Gamma_C \backslash \Gamma_C^\circ| \geq \gamma > 0$ uniformly in h. Moreover, assume that $\mathbf{u}_t \in H^1(\Gamma_C)$ holds and that the subsets

$$\Gamma_C^+ = \{\mathbf{x} \in \Gamma_C : \mathbf{u}_t > 0\}\,,\ \Gamma_C^- = \{\mathbf{x} \in \Gamma_C : \mathbf{u}_t < 0\} \tag{4.70}$$

are separated by a segment of length at least ℓ, where $\mathbf{u}_t \equiv 0$ holds. Then,

$$\| |\mathbf{u}_t| - |(\mathbf{u}_h^R)_t| \|_{1/2,\Gamma_C} \leq \frac{\tilde{C}_I}{\ell^{1/2}} \|\mathbf{u}_t - (\mathbf{u}_h^R)_t\|_{1/2,\Gamma_C} \tag{4.71}$$

holds with a constant $\tilde{C}_I$ independent of ℓ and h.

The proof will be based on real interpolation between $L^2(\Gamma_C)$ and $H^1(\Gamma_C)$ and the crucial part consists in showing the analogous estimate to (4.71) in $H^1(\Gamma_C)$ which is the content of the following lemma.

Lemma 4.16. *Let $(\mathbf{u}_h^R)_t$ satisfy the assumptions from Theorem 4.15. Then there exists a constant $C_I \geq 0$, depending only on the ratio M/m in (4.69), such that*

$$\| |\psi| - |(\mathbf{u}_h^R)_t| \|_{1,\Gamma_C} \leq \frac{C_I}{\ell} \|\psi - (\mathbf{u}_h^R)_t\|_{1,\Gamma_C} \tag{4.72}$$

holds for all $\psi \in H^1(\Gamma_C)$ with the property that the subsets

$$\Gamma_C^+ = \{\mathbf{x} \in \Gamma_C : \psi > 0\}\,,\ \Gamma_C^- = \{\mathbf{x} \in \Gamma_C : \psi < 0\} \tag{4.73}$$

are separated by a segment of length at least ℓ, where $\psi \equiv 0$ holds.

Proof. Without loss of generality, we may restrict ourselves to a subset of Γ_C, parametrized by $\mathbf{x}(s), s \in I = [0,1]$ such that the function $\phi = ((\mathbf{u}_h^R)_t) \circ \mathbf{x} \in H^1(I)$ vanishes on $[0,\xi]$ with $0 \leq \xi \leq 1$ and $m \leq \phi'(s) \leq M$ for $\xi < s < 1$. We define $I_- = \{s \in I : \psi(s) \leq 0\}$ and note that for all $s \in I \backslash I_-$ we have $|\psi(s)| - |\phi(s)| = \psi(s) - \phi(s)$ and therefore

$$\| |\psi| - |\phi| \|_{1,I\backslash I_-} = \|\psi - \phi\|_{1,I\backslash I_-}\,.$$

Consequently, if I_- is empty the assertion trivially holds. Otherwise I_- is the union of a number of subintervals $I_-^{(i)} = [\xi_-^{(i)}, \xi_+^{(i)}]$, $i = 1,2,3,\ldots N$, where $\psi \leq 0$ holds and which are at least of length ℓ due to our assumptions. We are left with showing

$$\| |\psi| - |\phi| \|_{1,I_-^{(i)}} \leq C_I \|\psi - \phi\|_{1,I_-^{(i)}} \text{ for } i = 1,2,3,\ldots N\,.$$

For each i, we have $I_-^{(i)} = I_l^{(i)} \cup I_r^{(i)}$ such that the following holds:

$$\phi \equiv 0 \text{ on } I_l^{(i)} \text{ , } m \leq \phi' \leq M \text{ on } I_r^{(i)} \text{ and } \psi \leq 0 \text{ on } I_-^{(i)} \text{ .} \tag{4.74}$$

Both the left and the right part of the interval $I_l^{(i)}$ may be empty. However, if the right part $I_r^{(i)}$ is not empty, it either is of length greater than ℓ or ψ vanishes on $I_r^{(i)}$ due to our assumptions on ψ and ϕ. In the latter case (i.e. $\psi \equiv 0$ on $I_r^{(i)}$) we have

$$\| |\psi| - |\phi| \|_{1,I_-^{(i)}} = \|\psi\|_{1,I_l^{(i)}} + \|\phi\|_{1,I_r^{(i)}} = \|\psi - \phi\|_{1,I_-^{(i)}} \text{ .}$$

We now consider the case of $|I_r^{(i)}| \geq \ell$. (4.74) implies

$$|\phi| = \phi \text{ and } |\psi| = -\psi \text{ on } I_-^{(i)}$$

and consequently

$$\frac{\| |\psi| - |\phi| \|^2_{1,I_-^{(i)}}}{\|\psi - \phi\|^2_{1,I_-^{(i)}}} = \frac{\|\psi + \phi\|^2_{1,I_-^{(i)}}}{\|\psi - \phi\|^2_{1,I_-^{(i)}}} = \frac{\|\psi + \phi\|^2_{0,I_-^{(i)}} + \|\psi' + \phi'\|^2_{0,I_-^{(i)}}}{\|\psi - \phi\|^2_{0,I_-^{(i)}} + \|\psi' - \phi'\|^2_{0,I_-^{(i)}}} \text{ ,} \tag{4.75}$$

which we need to bound from above by a constant. The first terms in the numerator and the denominator of (4.75) are related by

$$\|\psi + \phi\|^2_{0,I_-^{(i)}} \leq \|\psi - \phi\|^2_{0,I_-^{(i)}} \text{ ,} \tag{4.76}$$

since

$$|\psi + \phi| \leq |\psi| + |\phi| = -\psi + \phi = |\psi - \phi|$$

holds pointwise on $I_-^{(i)}$. The second term in the numerator may be bounded as

$$\begin{aligned} \|\psi' + \phi'\|^2_{0,I_-^{(i)}} &= \|\psi' - \phi' + 2\phi'\|^2_{0,I_-^{(i)}} \leq 2\|\psi' - \phi'\|^2_{0,I_-^{(i)}} + 8\|\phi'\|^2_{0,I_-^{(i)}} \\ &\leq 2\|\psi' - \phi'\|^2_{0,I_-^{(i)}} + 8M^2\|1\|^2_{0,I_r^{(i)}} \leq 2\|\psi' - \phi'\|^2_{0,I_-^{(i)}} + 24\frac{M^2}{m^2\ell^2}\|\phi\|^2_{0,I_r^{(i)}} \\ &\leq 2\|\psi' - \phi'\|^2_{0,I_-^{(i)}} + 24\frac{M^2}{m^2\ell^2}\|\psi - \phi\|^2_{0,I_-^{(i)}} \text{ ,} \end{aligned} \tag{4.77}$$

where we used the fact that ϕ lies above the linearly increasing function with slope m on $I_r^{(i)}$. Inserting (4.76) and (4.77) into (4.75) leads to

$$\frac{\| |\psi| - |\phi| \|^2_{1,I_-^{(i)}}}{\|\psi - \phi\|^2_{1,I_-^{(i)}}} \leq \max\left\{1 + 24\frac{M^2}{m^2\ell^2}, 2\right\} = 1 + 24\frac{M^2}{m^2\ell^2} \text{ ,}$$

which completes the proof. □

Before turning to the proof of Theorem 4.15 let us look at an example that (4.72) does not hold without the additional assumption on ψ.

Example 4.17. Let Γ_C be the interval $[0,1]$ and consider $(\mathbf{u}_h^R)_t$ to be the piecewise linear function which vanishes on $[0,1/2]$ and is monotonically increasing from 0 to $1/2$ on $[1/2,1]$. With ψ being the function defined by

$$\psi(s) = \begin{cases} -2\delta s & , \, s \in [0,1/2] \, , \\ s - 1/2 - \delta & , \, s \in [1/2,1] \end{cases}$$

we get

$$\|\psi' - ((\mathbf{u}_h^R)_t)'\|_{0,\Gamma_C}^2 = 2\delta^2 \, , \; \|\psi - (\mathbf{u}_h^R)_t\|_{0,\Gamma_C}^2 = \frac{2}{3}\delta^2 \, ,$$
$$\||\psi|' - |(\mathbf{u}_h^R)_t|'\|_{0,\Gamma_C}^2 = 2\delta^2 + 4\delta \, , \; \||\psi| - |(\mathbf{u}_h^R)_t|\|_{0,\Gamma_C}^2 = \frac{2}{3}\delta^2 - \frac{2}{3}\delta^3 \, .$$

Therefore, the ratio

$$\frac{\||\psi| - |(\mathbf{u}_h^R)_t|\|_{1,\Gamma_C}^2}{\|\psi - (\mathbf{u}_h^R)_t\|_{1,\Gamma_C}^2} = \frac{3}{2\delta} + 1 - \frac{\delta}{4}$$

becomes arbitrarily large as $\delta \to 0$.

Proof of Theorem 4.15. (i) As already indicated, the proof uses the interpretation of $H^{1/2}(\Gamma_C)$ as interpolation space $[L^2(\Gamma_C), H^1(\Gamma_C)]_{1/2,2}$ (cf. [AF03, Theorem 7.23]). The norm in $H^{1/2}(\Gamma_C)$ is therefore given by

$$\|\chi\|_{1/2,\Gamma_C} = \left(\int_0^\infty t^{-2} K(t,\chi)^2 \, dt \right)^{1/2}$$

with

$$K(t,\chi) = \inf_{\vartheta \in H^1(\Gamma_C)} \left(\|\chi - \vartheta\|_{0,\Gamma_C} + t \|\vartheta\|_{1,\Gamma_C} \right) \, .$$

We can therefore prove (4.71) by showing that

$$K(t, |\mathbf{u}_t| - |(\mathbf{u}_h^R)_t|) \leq C_I \, K(t, \mathbf{u}_t - (\mathbf{u}_h^R)_t)$$

holds for all $t \in [0,\infty)$. Equivalently, we will use

$$\hat{K}(t,\chi) = \inf_{\vartheta \in H^1(\Gamma_C)} \left(\|\chi - \vartheta\|_{0,\Gamma_C}^2 + t^2 \|\vartheta\|_{1,\Gamma_C}^2 \right)^{1/2} \, . \tag{4.78}$$

and show that

$$\hat{K}(t, |\mathbf{u}_t| - |(\mathbf{u}_h^R)_t|) \leq C_I \, \hat{K}(t, \mathbf{u}_t - (\mathbf{u}_h^R)_t)$$

is satisfied for all $t \in [0,\infty)$.

(ii) For $\chi \in H^1(\Gamma_C)$, the infimum in (4.78) is attained, for $t > 0$, by the solution $\theta_\chi(t) \in H^1(\Gamma_C)$ of

$$(\theta_\chi(t), \rho)_{0,\Gamma_C} + t^2 (\theta_\chi(t), \rho)_{1,\Gamma_C} = (\chi, \rho)_{0,\Gamma_C} \quad \forall \, \rho \in H^1(\Gamma_C) \tag{4.79}$$

which leads to

$$\hat{K}(t,\chi)^2 = \|\chi - \theta_\chi(t)\|^2_{0,\Gamma_C} + t^2\|\theta_\chi(t)\|^2_{1,\Gamma_C} =: Q_\chi(t)\,. \tag{4.80}$$

We have $\theta_\chi(0) = \chi$ and $\theta_\chi \to 0$ for $t \to \infty$ which leads to $Q_\chi(0) = 0$ and $Q_\chi(t) \to \|\chi\|^2_{0,\Gamma_C}$ for $t \to \infty$. Moreover, from (4.78) and (4.80), we deduce that $Q_\chi(t)$ is monotonically increasing in $[0,\infty)$. Differentiating (4.79) with respect to t gives

$$(\theta'_\chi(t),\rho)_{0,\Gamma_C} + 2t(\theta_\chi(t),\rho)_{1,\Gamma_C} + t^2(\theta'_\chi(t),\rho)_{1,\Gamma_C} = 0 \quad \forall\,\rho \in H^1(\Gamma_C)$$

and, in particular, $\theta'_\chi(0) = 0$. Differentiating (4.80) with respect to t implies

$$Q'_\chi(t) = 2(\chi - \theta_\chi(t), \theta'_\chi(t))_{0,\Gamma_C} + 2t\|\theta_\chi(t)\|^2_{1,\Gamma_C} + 2t^2(\theta_\chi(t),\theta'_\chi(t))_{1,\Gamma_C} \tag{4.81}$$

leading to

$$Q'_\chi(0) = 2(\chi - \theta_\chi(0), \theta'_\chi(0))_{0,\Gamma_C} = 0\,. \tag{4.82}$$

In order to access the behavior of $Q_\chi(t)$ near 0, the second derivative $Q''_\chi(0)$ is therefore needed. Differentiating (4.81) once more, we obtain

$$\begin{aligned} Q''_\chi(t) = {} & 2(\chi - \theta_\chi(t), \theta''_\chi(t))_{0,\Gamma_C} - 2\|\theta'_\chi(t)\|^2_{0,\Gamma_C} + 2\|\theta_\chi(t)\|^2_{1,\Gamma_C} \\ & + 8t(\theta_\chi(t),\theta'_\chi(t))_{1,\Gamma_C} + 2t^2\|\theta'_\chi(t)\|^2_{1,\Gamma_C} + 2t^2(\theta_\chi(t),\theta''_\chi(t))_{1,\Gamma_C} \end{aligned} \tag{4.83}$$

and, in particular,

$$Q''_\chi(0) = 2(\chi - \theta_\chi(0), \theta''_\chi(0))_{0,\Gamma_C} - 2\|\theta'_\chi(0)\|^2_{0,\Gamma_C} + 2\|\theta_\chi(0)\|^2_{1,\Gamma_C} = 2\|\chi\|^2_{1,\Gamma_C}\,.$$

The pointwise inverse triangle inequality gives us

$$\begin{aligned} \lim_{t\to\infty} Q_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(t) &= \||\mathbf{u}_t| - |(\mathbf{u}_h^R)_t|\|^2_{0,\Gamma_C} \\ &\le \|\mathbf{u}_t - (\mathbf{u}_h^R)_t\|^2_{0,\Gamma_C} = \lim_{t\to\infty} Q_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(t) \end{aligned}$$

and from Lemma 4.16 we deduce

$$\begin{aligned} Q''_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(0) &= 2\||\mathbf{u}_t| - |(\mathbf{u}_h^R)_t|\|^2_{1,\Gamma_C} \\ &\le 2\frac{C_I^2}{\ell^2}\|\mathbf{u}_t - (\mathbf{u}_h^R)_t\|^2_{1,\Gamma_C} = \frac{C_I^2}{\ell^2}Q''_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(0)\,. \end{aligned}$$

If we define $\tilde{Q}_\chi(s) = Q_\chi(s(\ell/C_I))$, then $\tilde{Q}''_\chi(0) = (\ell/C_I)^2 Q''_\chi(0)$ and therefore

$$\begin{aligned} \tilde{Q}''_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(0) &\le Q''_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(0)\,, \\ \lim_{s\to\infty} \tilde{Q}_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(s) &\le \lim_{t\to\infty} Q_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(t)\,. \end{aligned}$$

This implies the existence of a constant C (independent of ℓ) such that

$$\tilde{Q}_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(t) \le C Q_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(t) \text{ for all } t \in [0,\infty)$$

holds leading to

$$\begin{aligned}
\||\mathbf{u}_t| - |(\mathbf{u}_h^R)_t|\|_{1/2,\Gamma_C}^2 &= \int_0^\infty t^{-2} Q_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(t)\, dt \\
&= \int_0^\infty \left(\frac{\ell}{C_I} s\right)^{-2} Q_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(\frac{\ell}{C_I} s)\, \frac{\ell}{C_I} ds \\
&= \frac{C_I}{\ell} \int_0^\infty s^{-2} \tilde{Q}_{|\mathbf{u}_t|-|(\mathbf{u}_h^R)_t|}(s)\, ds \\
&\leq \frac{C_I C}{\ell} \int_0^\infty s^{-2} Q_{\mathbf{u}_t-(\mathbf{u}_h^R)_t}(s)\, ds = \frac{C_I C}{\ell} \|\mathbf{u}_t - (\mathbf{u}_h^R)_t\|_{1/2,\Gamma_C}^2
\end{aligned}$$

which completes the proof. □

The assumptions on $(\mathbf{u}_h^R)_t$ can be checked numerically and they do indeed hold in our numerical examples in section 4.6. The restriction to situations for which the solution satisfies $\mathbf{u}_t \in H^1(\Gamma_C)$ is certainly an unpleasant limitation. However, this is justified, at least for the test examples in section 4.6, by the behavior of $(\mathbf{u}_h^R)_t$ in our numerical experiments. The assumption on $\mathbf{u}_t$ vanishing on a segment of length ℓ between sign changes is physically reasonable and corresponds to the continuous sign change of the shear stress $(\boldsymbol{\sigma} \cdot \mathbf{n})_t$. Such an assumption also occurs in the uniqueness study for contact with Coulomb friction in [Ren06] and is used for finite element error analysis in [HR07]. In order to deduce the assumption $\gamma_I(0) = 0$ from the simple upper bound

$$\gamma_I(\mu_F) \leq \tilde{C}_I \frac{\mu_F}{\ell^{1/2}} \tag{4.84}$$

obtained from (4.61) and Theorem 4.15, we would need to have $\ell \approx \mu_F^{2-\varepsilon}$ as $\mu_F \to 0$ with $\varepsilon > 0$. The actual dependence of ℓ on μ_F is expected to depend on the geometry of the domain and of the obstacle considered as well as on the other material parameters μ and λ. From the numerical evidence it seems that this dependence is not quite reached in the first two examples of section 4.6 when a compressible material is considered. From the symmetry in this special situation, we may however deduce that (4.61) is fulfilled with $\gamma_I(\mu_F) = \mu_F$ if the sticky zone around the lowest part of the half-disk is adequately resolved. We will see in section 4.6 that we obtain good results even for rather large friction coefficients μ_F.

We also want to remark that the three-dimensional situation, with Γ_C being a surface, appears to be considerably more complicated and is beyond the scope of this work.

4.6 Numerical Experiments

In this section we will first review the examples in section 3.5 and test the extension of our method to the problem with Coulomb friction. After that we will compare the perfomance of our method to an example from the literature and finally we will present the results of a three-dimensional test case.

Example 1: Hertzian Contact - half-disk on straight line

The setting of this example is the same as in 3.5, with the only difference that we consider the Coulomb friction law with different friction parameters. We will again cover both compressible ($\lambda = 1$) and incompressible materials ($\lambda = \infty$) with the shear modulus μ scaled to 1 as always. We have run tests with the friction parameters μ_F equal to $0.15, 0.4$ and 1.35. The Dörfler parameter was set to 0.8.

Tables 4.1-4.3 show the results of the compressible case for $\mu_F = 0.15, 0.4$ and 1.35 respectively. Table s 4.4-4.6 represent the incompressible case. The number of active constraints for the friction condition are denoted by A_t. Since the displacement was (as in section 3.5) reconstructed in a way such that η_4 and η_5 vanished, we did not list it in the Table . If one chooses not to impose boundary conditions on $\mathbf{u}_h^R$ the frictional terms in the error estimator will not vanish but the resulting refinement and convergence are comparable. In a sense the error is 'pushed" into the other terms of the estimator by imposing boundary conditions on $\mathbf{u}_h^R$.

Figure 4.1 depicts the initial and deformed configurations of the compressible case for $\mu_F = 0.4$ after 5 steps of adaptive refinement. The deformed configuration is again obtained using the reconstructed displacement $\mathbf{u}_h^R \in \mathcal{P}_2(\mathcal{T}_h)$. While on a macroscopic level the refinement is indistinguishable from the frictionless case (cf. Figure 3.3), zooming into the contact zone as in Figure 4.2 reveals that refinement actually concentrates separately at the transition points from contact to separation and from stick to slip.

Taking a look at the forces and displacements in the highly resolved contact zone shown in Figure s 4.3-4.5 for the compressible case and 4.6-4.8 for the incompressible case, we observe several qualitative differences: First of all we notice that in the incompressible case, the shear stress features noisy oscillations. We attribute this to round-off errors induced by the high refinement of the curved boundary since this is not observed in the examples with straight boundaries below. Second we observe that, as expected, higher friction coefficients result in larger sticky zones. Furthermore we observe significantly larger sticky zones whenever the material is incompressible. This corresponds to the observation in 3.5, where in the incompressible case the tangential displacement almost vanished around the turning point even without the presence of friction. A related feature which can be observed in the incompressible case is that the shear stress changes its sign twice in the sticky part. We suspect this to be due to the materials resistance to volume change, causing it to want to move away from the central point of contact at (0,-0.5) but being constrained by friction. Further away from the central point of contact the body tends to slide inward and the sign of the constraining frictional shear stress changes. It then increases until sliding occurs. This behavior can best be captured with an adequately high resolution in the contact zone which is provided by the adaptive refinement strategy.

A comparison of the reduction of η for uniform and adaptive refinement is given in Figure 4.9. As for the frictionless problem adaptive refinement outperforms uniform refinement while failing to recover the optimal rate $\eta \sim N_h^{-1}$. Again the optimal rate is observed when retreating to problems with straight boundaries as illustrated by examples two and three below. A representative stress distribution is depicted in Figure 4.10.

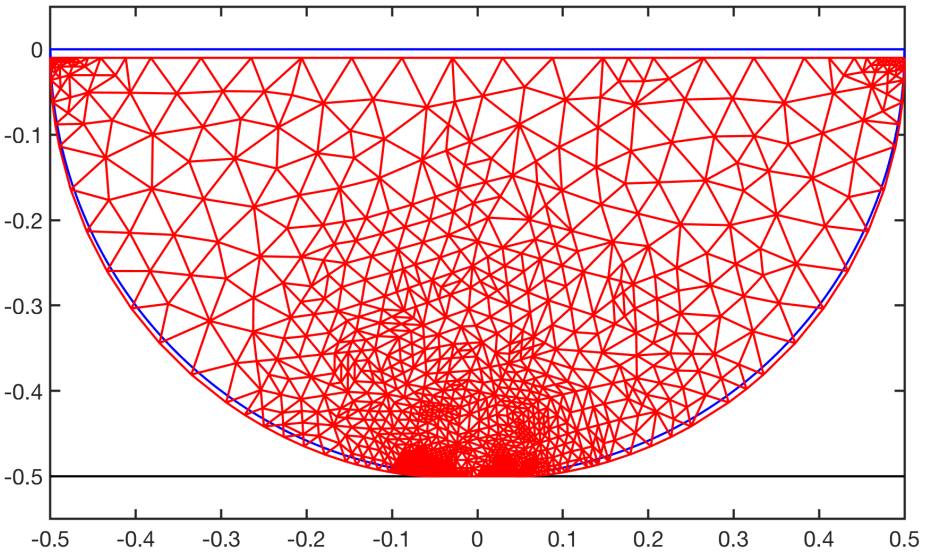

Figure 4.1: Ex. 1: Def. mesh after 5 refinements for $\lambda = 1$ and $\mu_F = 0.4$.

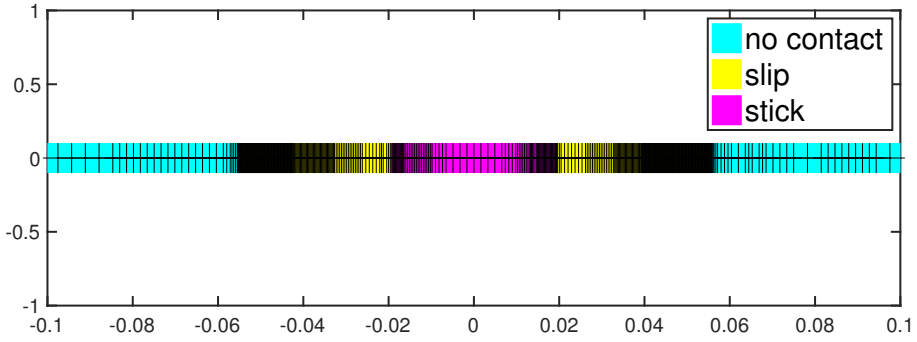

Figure 4.2: Ex. 1: Contact zone after 10 refinements for $\lambda = 1$ and $\mu_F = 0.4$.

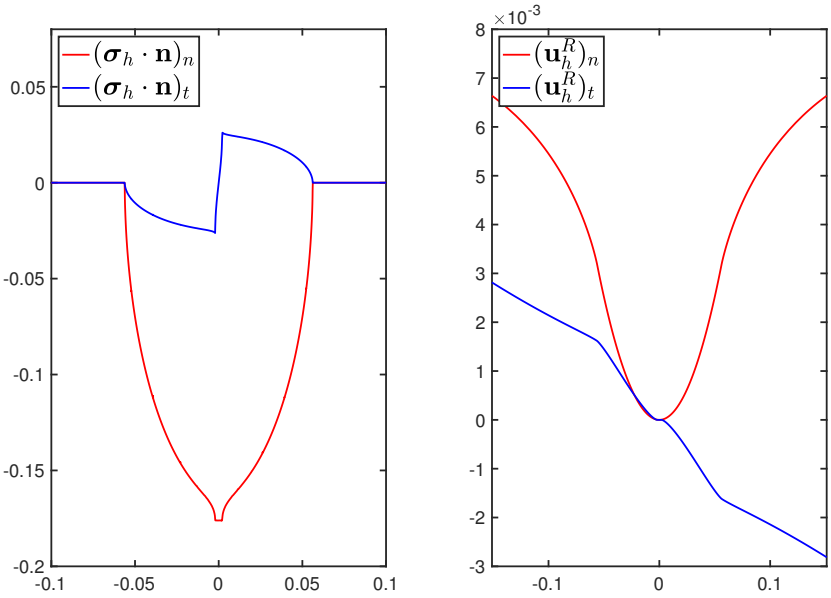

Figure 4.3: Ex. 1: Stress and displacement for $\lambda = 1$ and $\mu_F = 0.15$.

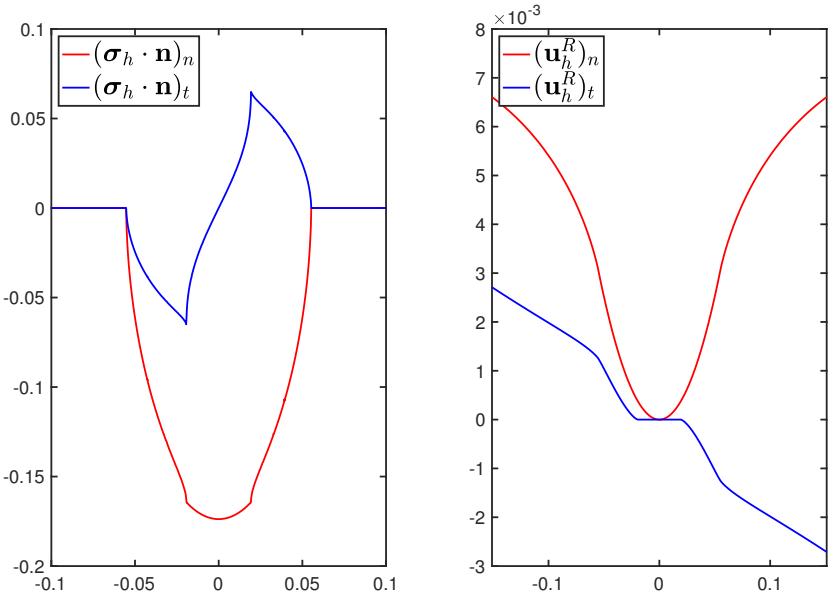

Figure 4.4: Ex. 1: Stress and displacement for $\lambda = 1$ and $\mu_F = 0.4$.

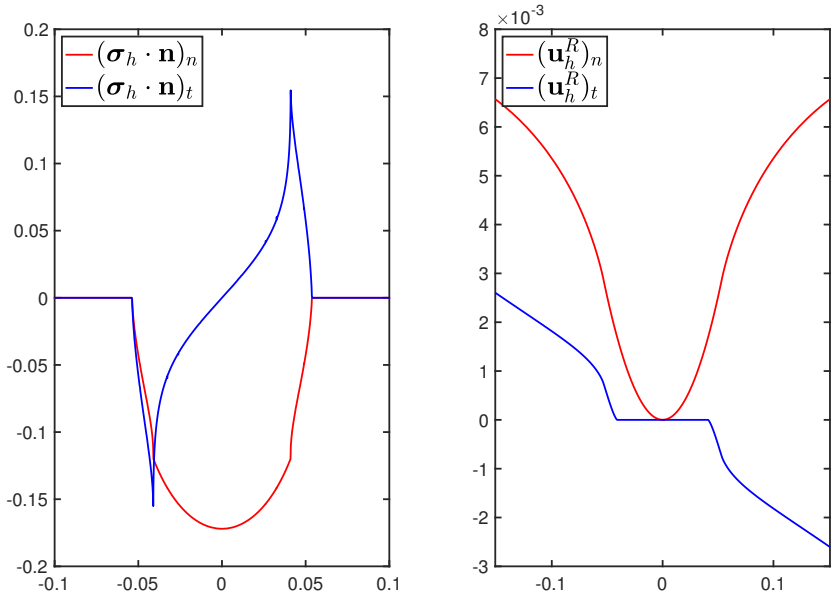

Figure 4.5: Ex. 1: Stress and displacement for $\lambda = 1$ and $\mu_F = 1.35$.

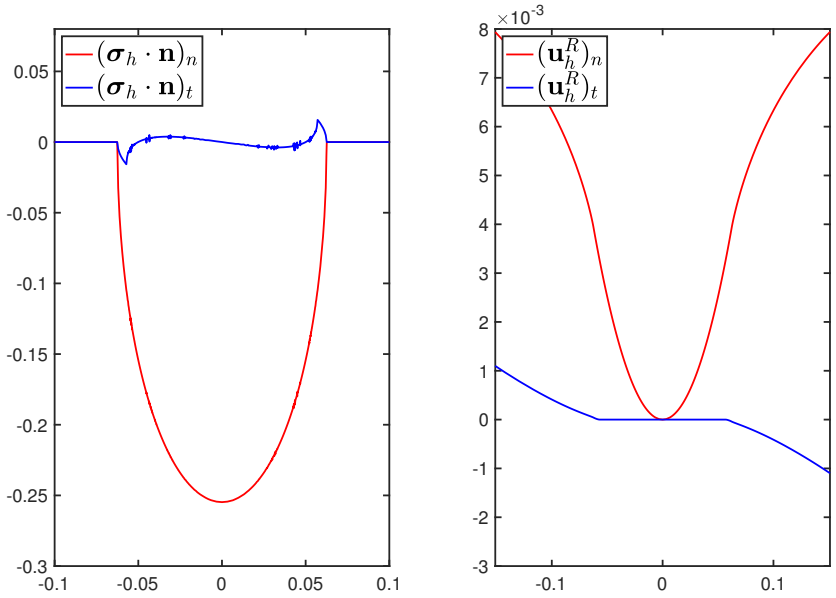

Figure 4.6: Ex. 1: Stress and displacement zone for $\lambda = \infty$ and $\mu_F = 0.15$.

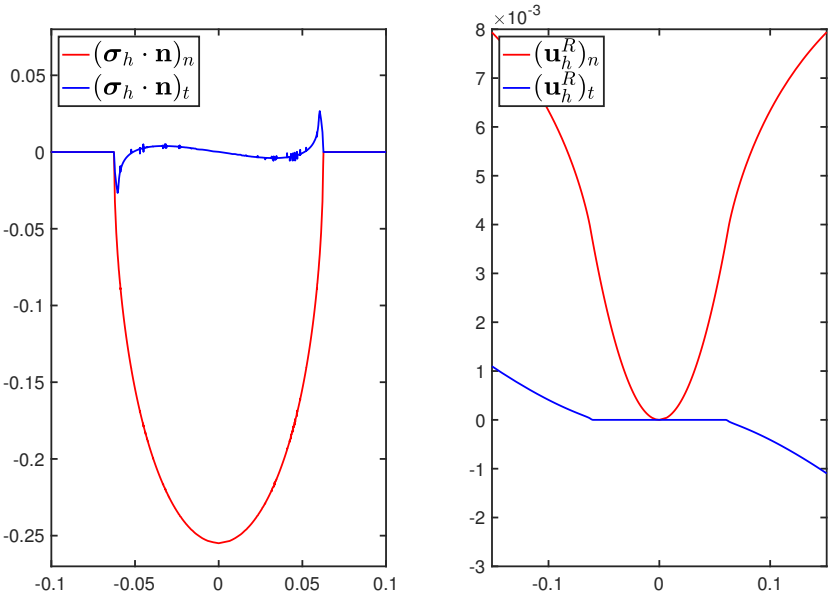

Figure 4.7: Ex. 1: Stress and displacement zone for $\lambda = \infty$ and $\mu_F = 0.4$.

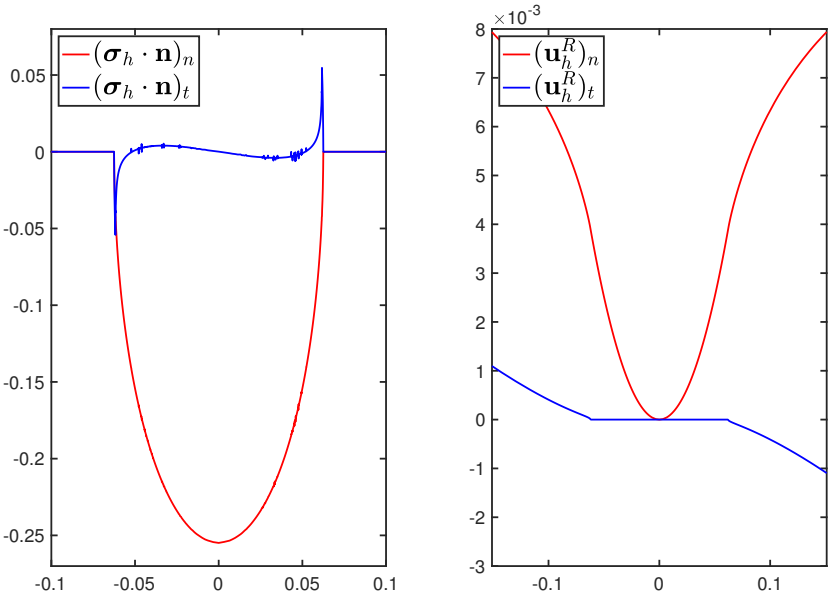

Figure 4.8: Ex. 1: Stress and displacement zone for $\lambda = \infty$ and $\mu_F = 1.35$.

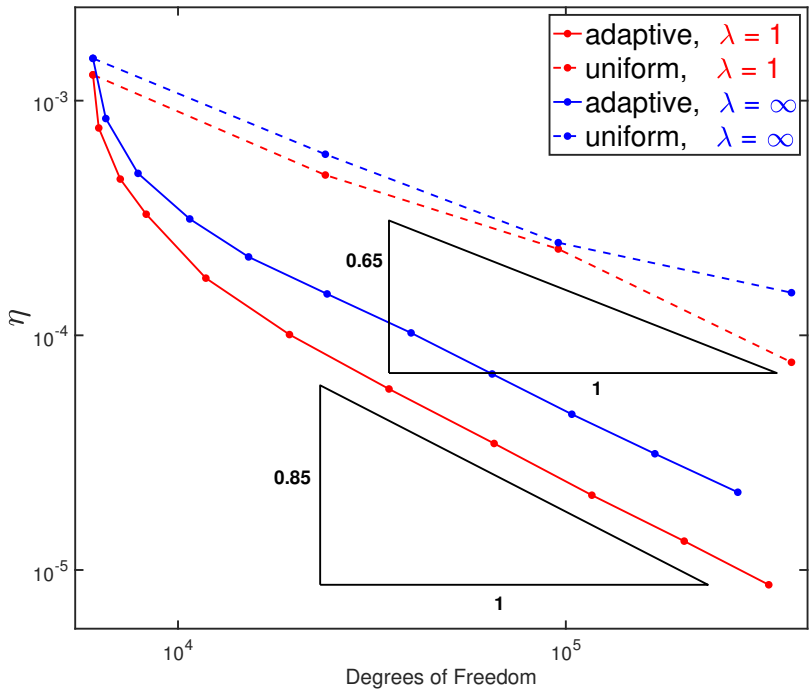

Figure 4.9: Example 1: Adaptive vs. uniform refinement for $\mu_F = 0.4$.

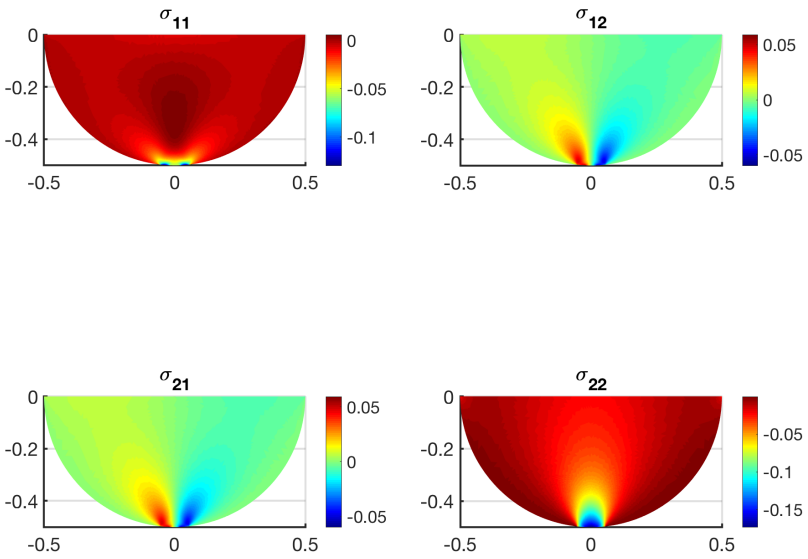

Figure 4.10: Example 1: Stress distribution for $\lambda = 1$ and $\mu_F = 0.4$.

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	20	0.000761618	0.000453669	0.000456915	0.00100016	
1	384	3864	2304	219	22	30	0.000427244	0.000304773	0.000341141	0.000643276	6.68
2	456	4596	2736	258	24	42	0.00024891	0.000187828	0.000177967	0.000359195	3.34
3	647	6524	3882	362	34	60	0.000123386	0.000103374	0.000107754	0.000193767	1.76
4	1037	10456	6222	567	44	84	8.7877e-05	7.57878e-05	7.04427e-05	0.000141467	0.67
5	1460	14740	8760	794	56	144	5.18342e-05	4.58174e-05	4.35809e-05	8.17638e-05	1.59
6	2986	30096	17916	1588	76	234	2.94699e-05	2.62525e-05	2.49092e-05	4.66705e-05	0.79
7	5348	53916	32088	2832	120	418	1.75235e-05	1.57965e-05	1.41856e-05	2.75288e-05	0.91
8	9797	98700	58782	5153	153	733	1.10512e-05	1.0028e-05	8.88807e-06	1.73692e-05	0.76
9	16649	167824	99894	8749	185	1333	7.13697e-06	6.47104e-06	5.73864e-06	1.12135e-05	0.82
10	27203	274580	163218	14353	257	2535	4.79743e-06	4.36603e-06	3.82754e-06	7.53177e-06	0.81

Table 4.1: Example 1: Results for Results for $\lambda = 1$ and $\mu_F = 0.15$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	18	0.000922237	0.000640837	0.000631607	0.00128846	
1	373	3748	2238	212	16	22	0.00052297	0.000373968	0.000414323	0.000764862	15.18
2	423	4256	2538	239	22	30	0.000311447	0.000243957	0.000240958	0.000463223	3.93
3	492	4956	2952	278	30	40	0.000214584	0.000178603	0.000172426	0.000328141	2.24
4	701	7068	4206	391	38	58	0.000111474	9.75328e-05	9.36846e-05	0.000175259	1.77
5	1154	11632	6924	628	44	82	6.35024e-05	5.65335e-05	5.40306e-05	0.000100737	1.12
6	2088	21036	12528	1115	74	138	3.70772e-05	3.32504e-05	3.17574e-05	5.90664e-05	0.9
7	3907	39328	23442	2062	96	216	2.16511e-05	1.94642e-05	1.87035e-05	3.46041e-05	0.86
8	6994	70320	41964	3655	122	339	1.30448e-05	1.17442e-05	1.12205e-05	2.08325e-05	0.87
9	12101	121692	72606	6301	162	589	8.32662e-06	7.48003e-06	7.13536e-06	1.32739e-05	0.82
10	19974	200936	119844	10399	226	1056	5.40846e-06	4.85378e-06	4.68853e-06	8.6483e-06	0.85

Table 4.2: Example 1: Results for $\lambda = 1$ and $\mu_F = 0.4$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	16	0.00108297	0.00101716	0.00070795	0.00164579	
1	378	3800	2268	215	22	22	0.000903929	0.000777229	0.000591175	0.00133066	4.38
2	400	4028	2400	228	26	28	0.000395263	0.00032643	0.000378876	0.000637445	12.35
3	466	4700	2796	264	30	34	0.000231082	0.000184903	0.000237541	0.000379491	3.34
4	619	6244	3714	348	34	44	0.000153008	0.000129792	0.000147888	0.000249255	1.48
5	861	8688	5166	477	46	62	8.53262e-05	7.54359e-05	7.88025e-05	0.000138495	1.78
6	1515	15280	9090	819	66	94	4.82412e-05	4.34666e-05	4.31549e-05	7.79673e-05	1.02
7	3005	30288	18030	1596	104	150	2.77733e-05	2.44715e-05	2.561e-05	4.5012e-05	0.81
8	5483	55204	32898	2889	130	218	1.65965e-05	1.42666e-05	1.53667e-05	2.67416e-05	0.87
9	10087	101524	60522	5275	176	342	1.02817e-05	8.57011e-06	9.74475e-06	1.65566e-05	0.79
10	17010	171236	102060	8887	237	540	6.64376e-06	5.3933e-06	6.44432e-06	1.07124e-05	0.83

Table 4.3: Example 1: Results for Results for $\lambda = 1$ and $\mu_F = 1.35$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	16	0.00111706	0.000762682	0.000681357	0.00151452	
1	388	3904	2328	221	20	22	0.000585331	0.000437907	0.00040097	0.000833757	7.83
2	469	4724	2814	268	24	25	0.000318919	0.000247357	0.000277659	0.000492257	2.75
3	637	6424	3822	359	32	36	0.000206716	0.000159565	0.000174256	0.000313939	1.46
4	911	9200	5466	507	40	49	0.000139411	0.000110262	0.000121154	0.000215168	1.05
5	1469	14836	8814	803	58	73	9.47579e-05	7.59001e-05	8.06925e-05	0.000145778	0.82
6	2403	24268	14418	1305	70	101	6.26936e-05	5.0674e-05	5.5152e-05	9.76769e-05	0.81
7	3862	38976	23172	2074	84	127	4.16776e-05	3.30373e-05	3.692e-05	6.47427e-05	0.87
8	6278	63380	37668	3357	116	183	2.7921e-05	2.21594e-05	2.47604e-05	4.34019e-05	0.82
9	10163	102656	60978	5431	140	265	1.89403e-05	1.4926e-05	1.69875e-05	2.94999e-05	0.8
10	16444	166176	98664	8777	164	380	1.284e-05	9.98508e-06	1.17406e-05	2.00611e-05	0.8

Table 4.4: Example 1: Results for $\lambda = \infty$ and $\mu_F = 0.15$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	16	0.00111706	0.000762682	0.000681357	0.00151452	
1	388	3904	2328	221	20	21	0.000579769	0.00044105	0.000416553	0.000839151	7.75
2	469	4724	2814	268	24	24	0.00031952	0.000246636	0.000277695	0.000489936	2.81
3	637	6424	3822	359	32	34	0.000205692	0.000158461	0.000175314	0.000313295	1.46
4	903	9116	5418	502	38	40	0.000139565	0.000109308	0.000123111	0.000215831	1.07
5	1440	14540	8640	787	52	56	9.68507e-05	7.64095e-05	8.5504e-05	0.000150098	0.78
6	2371	23948	14226	1288	64	72	6.57849e-05	5.24924e-05	5.83395e-05	0.000102404	0.77
7	3837	38748	23022	2066	82	98	4.37789e-05	3.42971e-05	3.98989e-05	6.84457e-05	0.84
8	6149	62144	36894	3304	112	142	2.94322e-05	2.31308e-05	2.6907e-05	4.61007e-05	0.84
9	10060	101708	60360	5401	126	172	1.99251e-05	1.54802e-05	1.8492e-05	3.12826e-05	0.79
10	16429	166244	98574	8825	170	243	1.36269e-05	1.04133e-05	1.28706e-05	2.14425e-05	0.77

Table 4.5: Example 1: Results for Results for $\lambda = \infty$ and $\mu_F = 0.4$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	361	3624	2166	205	16	16	0.00111706	0.000762682	0.000681357	0.00151452	
1	388	3904	2328	221	20	20	0.0005795	0.000445109	0.000424253	0.000844945	7.66
2	467	4704	2802	267	24	24	0.000318284	0.00024233	0.000278216	0.000487271	2.94
3	634	6392	3804	357	32	32	0.000209408	0.000159521	0.000180041	0.000318925	1.38
4	885	8940	5310	494	40	40	0.000143704	0.000113104	0.00012755	0.000222963	1.07
5	1386	14008	8316	762	58	60	0.000100209	7.92113e-05	8.85404e-05	0.000155421	0.8
6	2303	23272	13818	1254	68	71	6.66446e-05	5.30633e-05	5.99431e-05	0.000104165	0.79
7	3755	37944	22530	2029	86	94	4.45936e-05	3.51237e-05	4.09875e-05	7.0016e-05	0.81
8	6184	62508	37104	3325	112	121	2.99832e-05	2.33558e-05	2.78647e-05	4.71267e-05	0.79
9	10100	102152	60600	5431	128	148	2.03727e-05	1.55872e-05	1.91798e-05	3.20293e-05	0.79
10	16524	167308	99144	8901	164	191	1.39381e-05	1.06424e-05	1.32846e-05	2.20003e-05	0.76

Table 4.6: Example 1: Results for $\lambda = \infty$ and $\mu_F = 1.35$

Example 2: Rectangle on circular arc

Next we add friction to the the second example of section 3.5. We consider the same material parameters as well as Dirichlet data as in example 1, set the friction parameter to $\mu_F = 0.4$ and the Dörfler parameter to $\theta = 0.8$. Initial and deformed configuration of the incompressible case after 5 refinement steps are depicted in Figure 4.11, and the refinement of the contact zone is given in Figure 4.12.

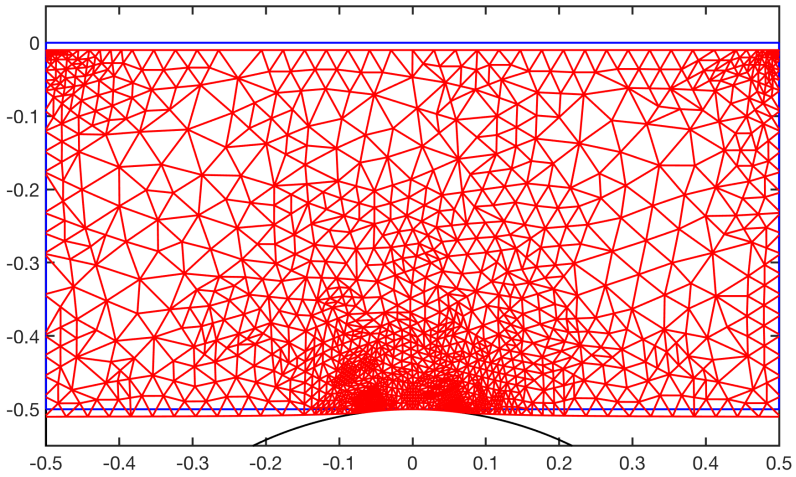

Figure 4.11: Ex. 2: Def. mesh after 5 refinements for $\lambda = \infty$ and $\mu_F = 0.4$.

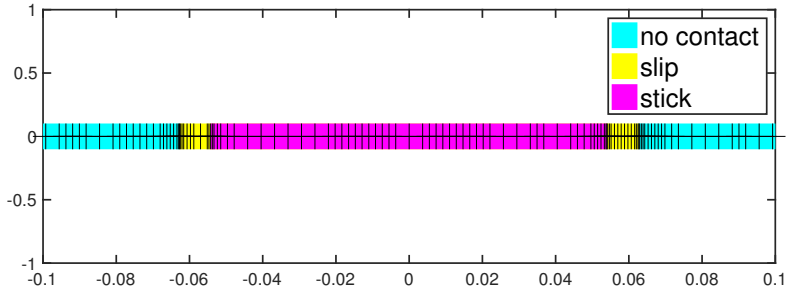

Figure 4.12: Ex. 2: Contact zone after 10 refinements for $\lambda = \infty$ and $\mu_F = 0.4$.

In Figure 4.13 we see that for this problem with straight boundaries there are no oscillations. Furthermore we see that the direction of the tangential displacement and correspondingly shear stress is reversed with respect to example 1.

Figure 4.14 presents again a comparison of the reduction of η for uniform and adaptive refinement, where now the optimal rate $\eta \sim N_h^{-1}$ is achieved, illustrating the efficiency of the proposed error estimator. The detailed results are summarized in Table s 4.7 and 4.8.

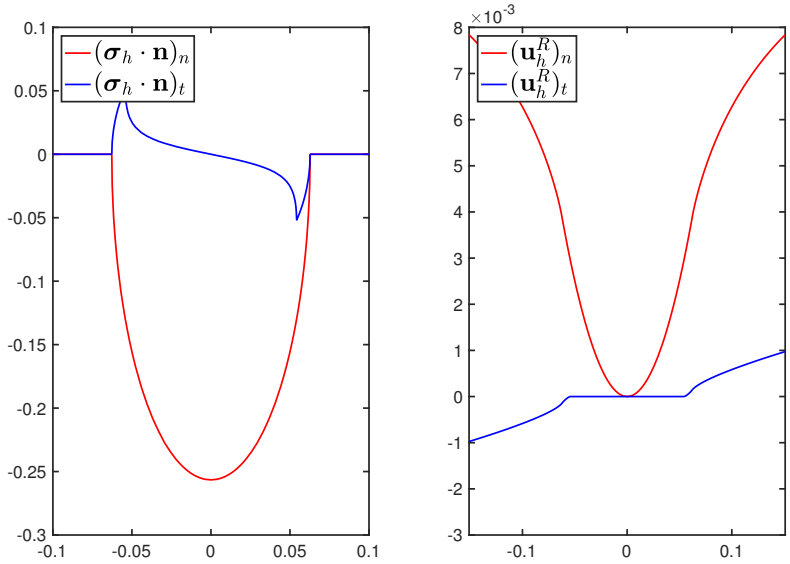

Figure 4.13: Ex. 2: Stress and displacement zone for $\lambda = \infty$ and $\mu_F = 0.4$.

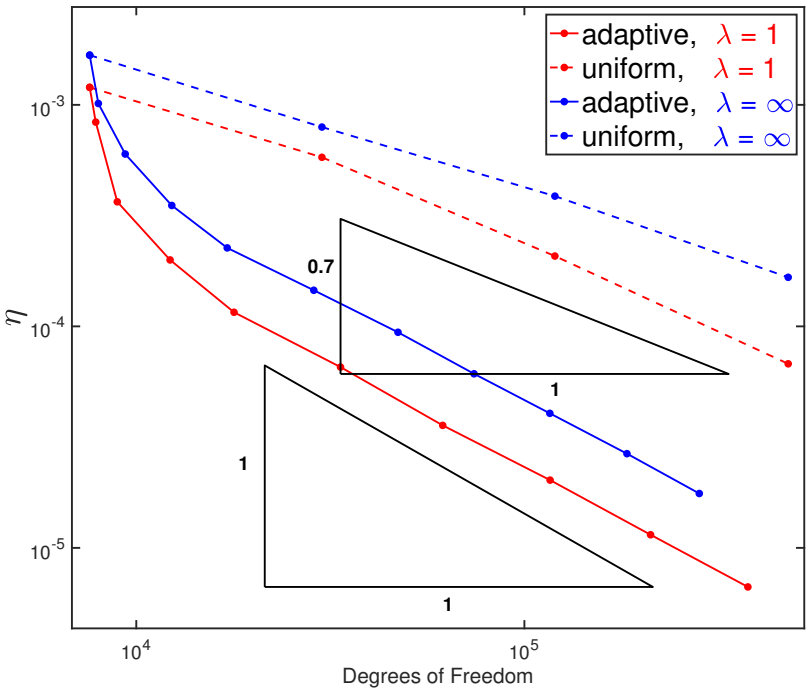

Figure 4.14: Example 2: Adaptive vs. uniform refinement for $\mu_F = 0.4$.

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	452	4552	2712	253	30	34	0.000902486	0.000530473	0.000527141	0.00119946	
1	468	4716	2808	262	32	36	0.000540378	0.000445223	0.000456582	0.000835883	10.07
2	531	5356	3186	296	38	46	0.00023101	0.0001879	0.000211074	0.000364999	6.48
3	726	7328	4356	401	48	62	0.000127633	0.000110162	0.000106586	0.000199465	1.93
4	1062	10716	6372	578	62	84	7.44777e-05	6.1551e-05	6.37881e-05	0.000115777	1.43
5	2004	20200	12024	1067	102	136	4.13929e-05	3.64297e-05	3.53336e-05	6.54901e-05	0.9
6	3693	37124	22158	1939	116	168	2.26707e-05	1.99842e-05	1.90571e-05	3.57282e-05	1
7	6997	70276	41982	3627	194	268	1.28336e-05	1.13963e-05	1.06972e-05	2.02239e-05	0.89
8	12733	127700	76398	6546	238	338	7.28792e-06	6.4847e-06	6.01593e-06	1.14611e-05	0.95
9	22714	227652	136284	11597	322	460	4.24466e-06	3.80869e-06	3.44518e-06	6.66277e-06	0.94

Table 4.7: Example 2: Results for $\lambda = 1$ and $\mu_F = 0.4$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	452	4552	2712	253	30	30	0.00114773	0.000755014	0.000959884	0.00167592	
1	475	4788	2850	266	32	32	0.000690381	0.000474968	0.000571664	0.00101441	9.79
2	556	5612	3336	312	38	42	0.000397521	0.000313561	0.00032166	0.00059984	3.3
3	733	7400	4398	407	48	52	0.000224628	0.0001911	0.000191048	0.000351392	1.93
4	1021	10300	6126	558	54	64	0.000143495	0.000119431	0.000127529	0.000226093	1.34
5	1709	17232	10254	921	86	99	9.27955e-05	7.93355e-05	7.92213e-05	0.000145537	0.86
6	2828	28452	16968	1503	106	122	6.01235e-05	5.17347e-05	5.05225e-05	9.40417e-05	0.87
7	4453	44732	26718	2332	122	144	3.89816e-05	3.30778e-05	3.34278e-05	6.1083e-05	0.96
8	6990	70192	41940	3636	176	204	2.58564e-05	2.21676e-05	2.1899e-05	4.0491e-05	0.91
9	11065	111004	66390	5716	220	256	1.6998e-05	1.45956e-05	1.43848e-05	2.66249e-05	0.92
10	17018	170592	102108	8734	264	309	1.11996e-05	9.51335e-06	9.68285e-06	1.75981e-05	0.96

Table 4.8: Example 2: Results for $\lambda = \infty$ and $\mu_F = 0.4$

Example 3: Unit square with volume force

Our third example is taken from [HL09] and features a volume force instead of prescribed displacement. Furthermore the tangential displacement of the solution is pointing only in one direction on all of Γ_C, and thus 4.61 holds without further requirements.

We consider the unit square $\Omega = (0,1) \times (0,1)$ which is clamped ($\mathbf{g}_D = \mathbf{0}$) on $\Gamma_D := (0,1)\times\{1\}$. Surface traction forces on $\Gamma_N := \{0,1\}\times(0,1)$ are also set to zero. The potential contact boundary is given by $\Gamma_C = (0,1) \times \{0\}$ and the obstactle is represented by the x_1 axis, resulting in a vanishing gap-function $g_C \equiv 0$. A uniform horizontal volume force $\mathbf{f} = (f_1, 0)^\mathsf{T}$ with $f_1 = 76518$ is acting on Ω. The friction parameter is set to 0.2 and the material parameters are given by the youngs modulus $E = 10^6$ and Poisson ratio $\nu = 0.3$. This corresponds to the Lamé parameters $\mu \approx 384.615$ and $\lambda \approx 576.923$ (cf. [Bra13, Chapter VI]).

As can be seen in Figure s 4.15 and 4.16 the error estimator captures both the corner singularities and the transition point from contact to separation.

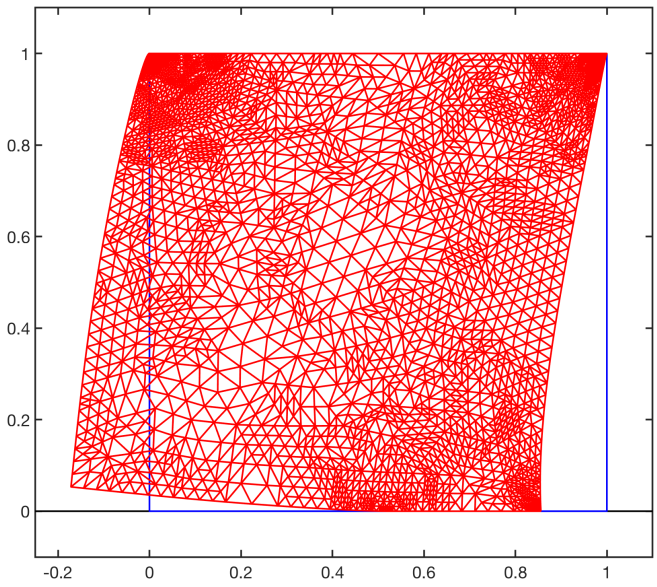

Figure 4.15: Example 3: Deformed mesh after 10 refinements.

The surface forces and displacements on Γ_C are plotted in Figure 4.17 and qualitatively resemble the results in [HL09]: There is no sticky zone and contact pressure peaks at $x_1 \approx 0.96$ with $\max(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n \approx 80800$. However, the transition between contact and separation at $x_1 \approx 0.67$ seems to be sharper in our computations. This is probably due to the adaptive refinement that is concentrated at the transition point. Morover, the tangential displacement appears to be shifted by 0.12 in our computations.

A comparison of the reduction of η for uniform and adaptive refinement is given in Figure 4.18 and the numerical details are summarized in Table 4.9.

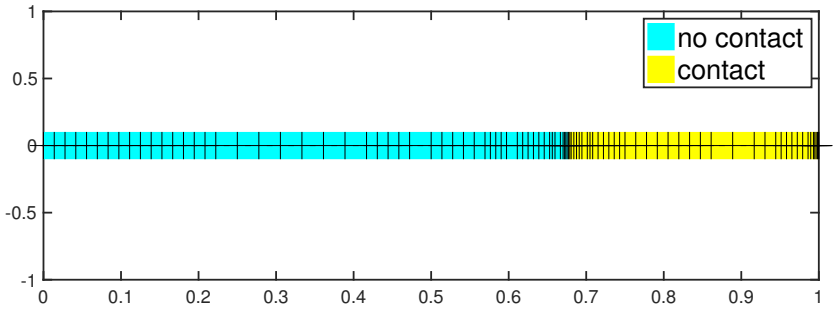

Figure 4.16: Example 3: Contact zone after 12 refinements.

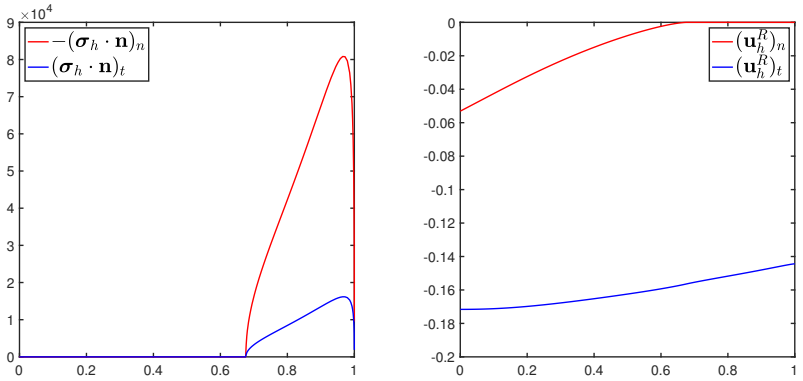

Figure 4.17: Example 3: Stress and displacement in contact zone.

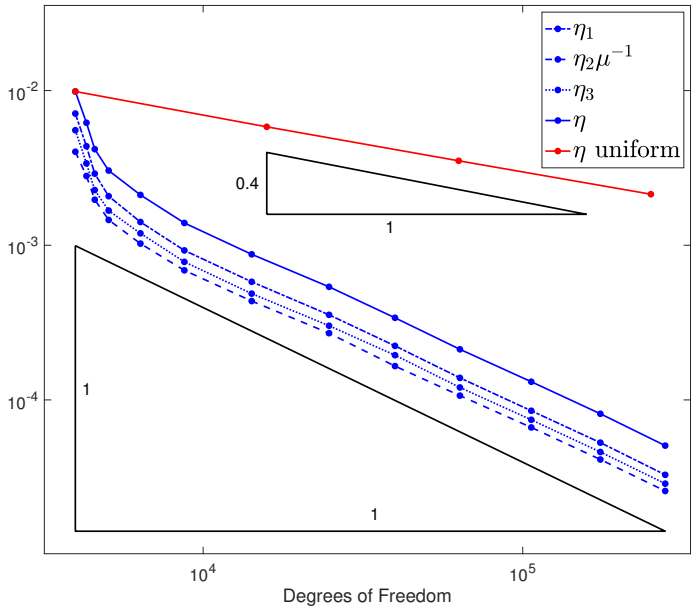

Figure 4.18: Example 3: Adaptive vs. uniform refinement.

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	$\mu^{-1}\eta_2$	η_3	η	eoc
0	238	2380	1428	138	12	18	0.00710049	0.00554089	0.00402267	0.00986409	
1	258	2580	1548	150	12	18	0.00436488	0.00338219	0.00280296	0.00619257	5.81
2	274	2740	1644	160	12	18	0.00290178	0.00227365	0.00196943	0.00417953	6.57
3	303	3032	1818	177	12	18	0.00207534	0.00167535	0.00145764	0.00303949	3.18
4	380	3812	2280	221	15	26	0.00141333	0.00119363	0.00102463	0.00211474	1.59
5	523	5240	3138	301	15	28	0.00092587	0.000780713	0.000688262	0.001393	1.32
6	852	8528	5112	477	19	36	0.000580194	0.000486515	0.00043553	0.000873503	0.96
7	1491	14924	8946	814	27	46	0.000354855	0.000301662	0.000270036	0.000538369	0.87
8	2401	24016	14406	1288	35	64	0.000223599	0.000194582	0.000165458	0.000339463	0.97
9	3838	38372	23028	2039	45	82	0.000138825	0.000120334	0.000106578	0.000212395	1
10	6422	64232	38532	3370	63	114	8.50605e-05	7.43771e-05	6.63741e-05	0.000131045	0.94
11	10575	105748	63450	5488	79	144	5.29159e-05	4.60405e-05	4.11834e-05	8.13382e-05	0.96
12	16883	168824	101298	8704	103	186	3.27801e-05	2.87399e-05	2.57643e-05	5.06391e-05	1.01

Table 4.9: Example 3: Results for $\mu_F = 0.2$

l	$\lvert\mathcal{T}_h\rvert$	$\dim \mathbf{\Sigma}_h$	$\dim \mathbf{U}_h$	$\dim \mathbf{\Theta}_h$	A_n	A_t	η_1	η_2	η_3	η	eoc
0	256	6944	3072	255	20	0	0.00612438	0.00129459	0.00549758	0.00833111	
1	304	8248	3648	294	52	16	0.00467775	0.00208991	0.00383390	0.00639905	1.5
2	728	19724	8736	573	164	60	0.00256599	0.00120344	0.00209277	0.00352310	0.68
3	2288	61932	27456	1560	516	216	0.00099197	0.00075432	0.00076786	0.00146377	0.77
4	8032	217852	96384	5130	1108	896	0.00059182	0.00036891	0.00045595	0.00083321	0.45

Table 4.10: Example 4: Results for $\mu_F = 0.4$

Example 4: Hertzian contact in 3D

For our forth example we consider the three-dimensional analogon to the setting examined in example 1. We consider the lower hemisphere with center at the origin and radius $R = 0.5$. The body is constrained by a rigid foundation represented by the horizontal plane at $x_3 = -0.5$ and the potential contact boundary is

$$\Gamma_C := \left\{(x_1, x_2, (R^2 - x_1^2 - x_2^2)^{1/2}) : |x_1| + |x_2| < 0.3\right\}.$$

The Dirichlet boundary consists of the top disk

$$\Gamma_D := \left\{(x_1, x_2) : x_1^2 + x_2^2 < R^2\right\} \times \{0\}$$

and displacement thereon is prescribed by $\mathbf{g}_D = (0, 0, -0.025)^\mathsf{T}$. Surface traction forces $\mathbf{g}_N$ on $\Gamma_N := \Gamma \setminus \left(\overline{\Gamma_D \cup \Gamma_C}\right)$ as well as volume forces $\mathbf{f}$ on Ω are set to zero. The Lamé parameters μ and λ are both set to 1 and the friction parameter is set to $\mu_F = 0.4$. The parameter for the Dörfler marking strategy is set to $\theta = 0.9$.

A plot of the reference and deformed configuration is given in Figure 4.19. The visualizations of the contact pressure in Figure 4.21 and of one component of the shear stress in Figure 4.22 mimic the results of the two-dimensional problem. The three-dimensional structure of the total contact force (both pressure and shear forces) can be seen in Figure 4.23, where it is represented as a vectorfield. The refinement is not yet small enough to detect two separate transition zones as in our two-dimensional examples. However, it can clearly be seen in Figure 4.20 that the refinement concentrates at the transition zone between contact and separation.

Since the computational cost increases significantly faster for simulations in 3D, we were not able to perform computations for uniformly refined meshes. Thus we can not present a comparison of the reduction of the error estimator for adaptive versus uniform refinement but have to restrict ourselves to reporting the reduction for our adaptive computations in Figure 4.24 and Table 4.10. The fact that adaptive refinement allowed for computations within a reasonable amount of time, is in and of itself a strong argument for the use of such a strategy.

The optimal rate one would hope to achieve is $\eta \sim N_h^{-2/3}$. It is hard to tell whether this is the case for our method, since our computations only seem to cover the pre-asymptotic regime. However, from what we have seen in our two-dimensional experiments we would expect the rate to be somewhere between 0.5 and 0.55 due to the curved boundary. While the results obtained so far are promising, further improvements to accelerate the solution method are necessary in order to perform computations on finer grids.

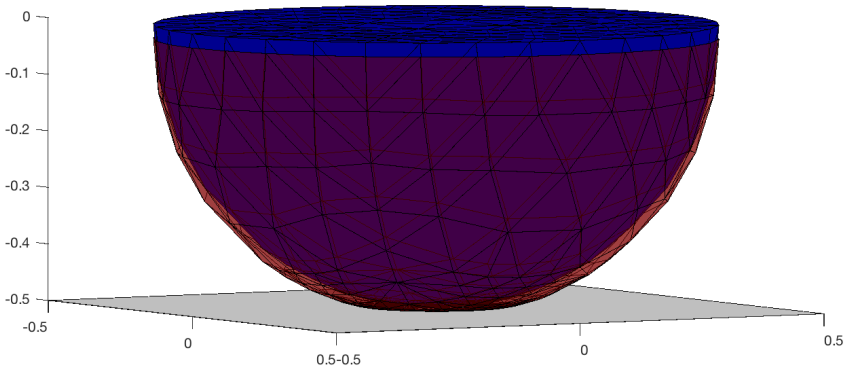

Figure 4.19: Example 4: Reference and deformed mesh after 4 refinements.

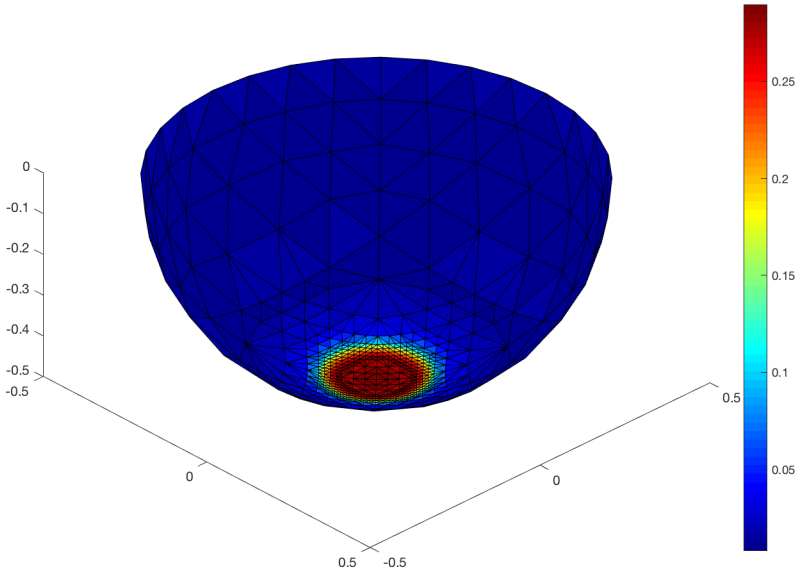

Figure 4.20: Example 4: Von Mises stress distribution (view from below).

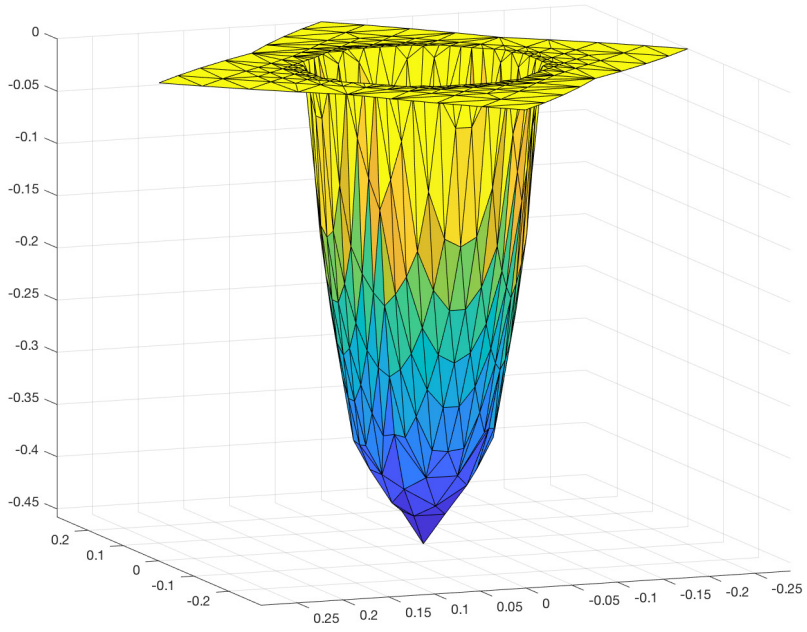

Figure 4.21: Example 4: Contact pressure $(\boldsymbol{\sigma}_h \cdot \mathbf{n})_n$ on Γ_C.

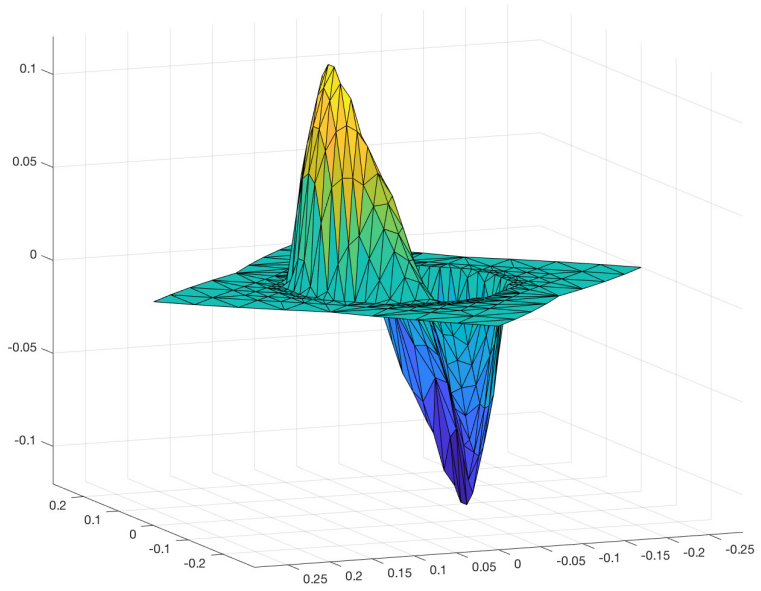

Figure 4.22: Example 4: One component of shear stress $[(\boldsymbol{\sigma}_h \cdot \mathbf{n})_t]_1$ on Γ_C.

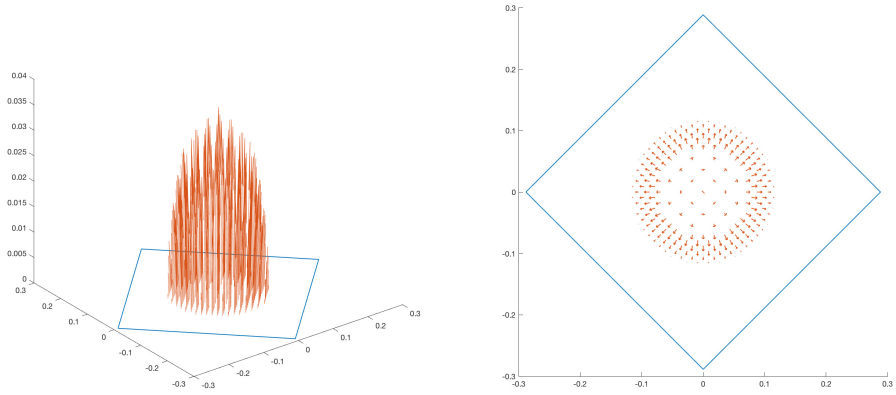

Figure 4.23: Example 4: Total contact force $\boldsymbol{\sigma}_h \cdot \mathbf{n}$ on Γ_C.

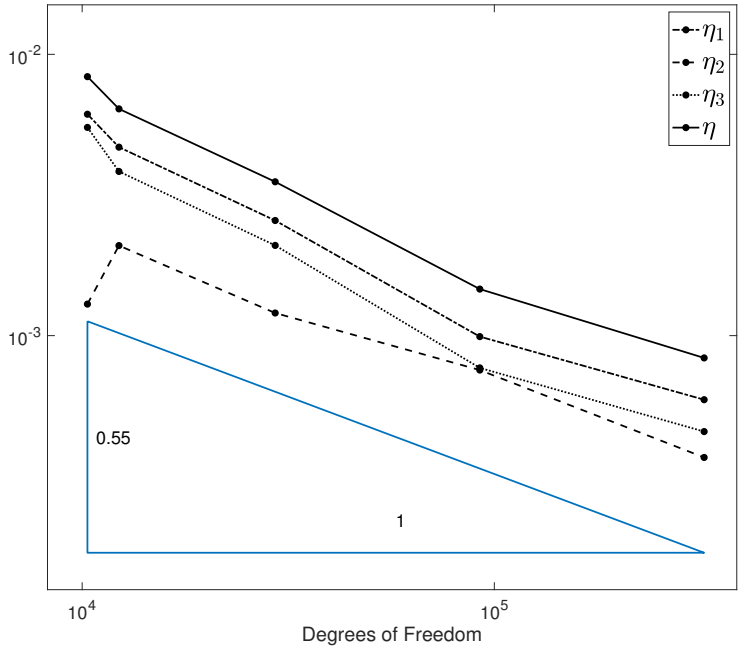

Figure 4.24: Example 4: Convergence behaviour for adaptive refinement.

Chapter 5

Conclusion and Outlook

Conclusion

We extended the finite element method for stress-based Linear Elasticity with weak symmetry proposed in [BBF09] to the Signorini problem with and without friction and adapted the displacement reconstruction procedure proposed in [Voh10] to respect the contact conditions. Exploiting the strong duality that holds for contact problems we derived a reliable a-posteriori error estimator and tested its efficiency on a number of numerical experiments in two and three dimensions. In the case of Coulomb friction the reliability result is obtained under somewhat particular but reasonable regularity assumptions on the solution. Finally we studied the application of a semi-smooth Newton method to solve the finite dimensional variational inequalities arising from the finite element discretization.

While the derivation of proper stress-based variational formulations for contact problems was achieved without regularizations of any kind, the numerical analysis of the corresponding finite element methods remained incomplete due to the lack of appropriate interpolation operators. The semi-smooth Newton method worked very well in 2D for both frictionless and the frictional problems. However, the convergence was slowed down significantly in 3D due to the stronger nonlinearity of the friction constraint formulated in terms of the Euclidean norm instead of the absolute value.

Even though optimal convergence rates were only recovered on domains with polygonal boundaries, the adaptive refinement procedure driven by the proposed a-posteriori error estimator achieved better rates than uniform refinement in all of our test cases, including compressible and incompressible materials as well as different friction parameters. Furthermore the higher resolution of the transition points in the contact zone and of corner singularities achieved by adaptive refinement caused a steep initial decrease of the error estimate, resulting in significantly smaller absolute error estimates relative to the number of degrees of freedom.

Outlook

During the course of this work there arose several questions that were beyond the scope of this thesis, but might be worth investigating in more detail:

As mentioned above the convergence of the semi-smooth Newton method was slowed down immensely when applied to 3D problems. In order to accelerate the solution process one could either try to speed up the Newton method or seek to develop alternative approaches such as the multilevel strategy proposed in [Kra09] for the primal problem. The former approach could be achieved by using iterative solvers for the solution of the linearized problem in each step and by replacing the Euclidean norm that appears in the friction constraint in 3D by a piecewise linear approximation. The latter approach is particularly challenging because of the saddlepoint structure inherent to mixed methods and the large kernel of the divergence operator (cf. [RKSK18]).

Another open question is the derivation of a priori estimates for the frictional problem along with a proof of convergence of the Tresca iteration introduced in section 4.3. The corresponding results for the primal formulation can be found in [HHN96]. However, the type of discretization of the friction condition that is needed to make the numerics work unfortunately does not allow the application of the methods and ideas used for the primal formulation. This remains a major drawback of the approach proposed in this work.

Furthermore, since the inequality result presented in section 4.5.1 is restricted to the setting in two dimensions, an extension to the three-dimensional case should be pursued. However, we expect this to be significantly more complex, since the dichotomy of positive and negative tangential displacement is lost when the contact boundary consists of a two-dimensional surface.

Proving that the proposed error estimator is also a lower bound appears to be a very challenging problem even for the frictionless case. We are however confident, that one would be able to recover the optimal rate also on domains with curved boundaries by incorporating the parametric Raviart-Thomas elements introduced in [BMS14].

Beside these open questions further research could focus on an extension of the concepts of this work to include additional phenomena in solid mechanics. Since plasticity is also a stress-driven process, the direct access to the stress-tensor-field provided by our approach is expected to be of advantage. The techniques in [Sta07] could be a starting point for the inclusion of deformations of elasto-plastic nature into our model. Furthermore the ideas in [Cap14] could be used to investigate alternative non-local friction laws as well as time dependent problems starting with a quasi-static formulation. For this purpose both adaptive time-stepping and space-time approaches could be considered.

Bibliography

[AAW08] D. N. Arnold, G. Awanou, and R. Winther. Finite elements for symmetric tensors in three dimensions. *Math. Comp.*, 77:1229–1251, 2008.

[ABD84] D. N. Arnold, F. Brezzi, and J. Douglas. PEERS: A new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1:347–367, 1984.

[ACS09] F. S. Attia, Z. Cai, and G. Starke. First-order system least squares for the Signorini contact problem in linear elasticity. *SIAM J. Numer. Anal.*, 47:3027–3043, 2009.

[AF03] R. A. Adams and J. F. Fournier. *Sobolev Spaces*. Academic Press, New York, 2nd edition, 2003.

[AH09] K. E. Atkinson and W. Han. *Theoretical Numerical Analysis*. Springer, New York, 3rd edition, 2009.

[AT79] M. Amara and J. M. Thomas. Equilibrium finite elements for the linear elasticity problem. *Numer. Math.*, 33:367–383, 1979.

[AW02] D. N. Arnold and R. Winther. Mixed finite elements for elasticity. *Numer. Math.*, 92:401–419, 2002.

[BBF09] D. Boffi, F. Brezzi, and M. Fortin. Reduced symmetry elements in linear elasticity. *Commun. Pure Appl. Anal.*, 8:95–121, 2009.

[BBF13] Daniele Boffi, Franco Brezzi, and Michel Fortin. *Mixed Finite Element Methods and Applications*. Springer, Heidelberg, 2013.

[BKMS18] F. Bertrand, B. Kober, M. Moldenhauer, and G. Starke. Weakly symmetric stress equilibration and a posteriori error estimation for linear elasticity. *Submitted for Publication, available at arXiv:1808.02655*, 2018.

[BMS14] F. Bertrand, S. Münzenmaier, and G. Starke. First-order system least squares on curved boundaries: Higher-order Raviart-Thomas elements. *SIAM J. Numer. Anal.*, 52:3165–3180, 2014.

[BMS19] Fleurianne Bertrand, Marcel Moldenhauer, and Gerhard Starke. A posteriori error estimation for planar linear elasticity by stress reconstruction. *Computational Methods in Applied Mathematics*, 19(3):663–679, 2019.

[Bra13] D. Braess. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer, Berlin, 2013. 5. Auflage.

[BS08] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 3rd edition, 2008.

[Cap14] A. Capatina. *Variational Inequalities and Frictional Contact Problems*. Springer, Cham, 2014.

[CD98] Carsten Carstensen and Georg Dolzmann. A posteriori error estimates for mixed fem in elasticity. *Numerische Mathematik*, 81(2):187–209, 1998.

[Cia88] P. G. Ciarlet. *Mathematical Elasticity Volume I: Three–Dimensional Elasticity*. North-Holland, Amsterdam, 1988.

[Cla13] Francis Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264. Springer Science & Business Media, 2013.

[Clé75] Ph Clément. Approximation by finite element functions using local regularization. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(R2):77–84, 1975.

[CS03] Z. Cai and G. Starke. First-order system least squares for the stress-displacement formulation: Linear elasticity. *SIAM J. Numer. Anal.*, 41:715–730, 2003.

[CS04] Z. Cai and G. Starke. Least squares methods for linear elasticity. *SIAM J. Numer. Anal.*, 42:826–842, 2004.

[DDE12] Françoise Demengel, Gilbert Demengel, and Reinie Erné. *Functional spaces for the theory of elliptic partial differential equations*. Springer, 2012.

[DE12] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer, Berlin, 2012.

[DL76] G. Duvaut and J. L. Lions. *Inequalities in Mechanics and Physics*. Springer, New York, 1976.

[EG04] Alexandre Ern and Jean-Luc Guermond. *Theory and Practice of Finite Elements*. Springer, New York, 2004.

[EJ98] Christof Eck and Jiří Jarušek. Existence results for the static contact problem with coulomb friction. *Mathematical Models and Methods in Applied Sciences*, 8(03):445–468, 1998.

[GG11] Jay Gopalakrishnan and Johnny Guzmán. Symmetric nonconforming mixed finite elements for linear elasticity. *SIAM J. Numer. Anal.*, 49:1504–1520, 2011.

[Gol60] Werner Goldsmith. *Impact.* Arnold, 1960.

[GR86] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer, New York, 1986.

[Gri85] Pierre Grisvard. *Elliptic Problems in Nonsmooth Domains.* SIAM, Boston, 1985.

[HHN96] J. Haslinger, I. Hlaváček, and J. Nečas. *Numerical Methods for Unilateral Problems in Solid Mechanics.* 1996. Handb. Numer. Anal. IV, P.G. Ciarlet and J. L. Lions eds., North-Holland, Amsterdam, pp. 313-485.

[HHNL88] I. Hlaváček, J. Haslinger, J. Nečas, and J. Lovíšek. *Solution of Variational Inequalities in Mechanics.* Springer, New York, 1988.

[HIK02] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.

[Hil03] Patrick Hild. An example of nonuniqueness for the continuous static unilateral contact model with coulomb friction. *Comptes Rendus Mathematique*, 337(10):685–688, 2003.

[HL09] Patrick Hild and Vanessa Lleras. Residual error estimators for coulomb friction. *SIAM Journal on Numerical Analysis*, 47(5):3550–3583, 2009.

[HR07] Patrick Hild and Yves Renard. An error estimate for the signorini problem with coulomb friction approximated by finite elements. *SIAM Journal on Numerical Analysis*, 45(5):2012–2031, 2007.

[Jar83] Jiří Jarušek. Contact problems with bounded friction. coercive case. *Czechoslovak Mathematical Journal*, 33(2):237–261, 1983.

[Kim11] Kwang-Yeon Kim. Guaranteed a posteriori error estimator for mixed finite element methods of linear elasticity with weak stress symmetry. *SIAM Journal on Numerical Analysis*, 49(6):2364–2385, 2011.

[KKS+08] Ralf Kornhuber, Rolf Krause, Oliver Sander, Peter Deuflhard, and Susanne Ertel. A monotone multigrid solver for two body contact problems in biomechanics. *Computing and Visualization in Science*, 11(1):3–15, 2008.

[KMS17] Rolf Krause, Benjamin Müller, and Gerhard Starke. An adaptive least-squares mixed finite element method for the signorini problem. *Numerical Methods for Partial Differential Equations*, 33(1):276–289, 2017.

[KO88] N. Kikuchi and J. T. Oden. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*. SIAM, Philadelphia, 1988.

[Kob17] Bernhard Kober. Symmetrie und Konformität in spannungsbasierter FEM für Lineare Elastizität. Master's thesis, Fakultät für Mathematik - Universität Duisburg-Essen, 2017.

[Kra06] R Krause. From inexact active set strategies to nonlinear multigrid methods. In *Analysis and Simulation of Contact Problems*, pages 13–21. Springer, 2006.

[Kra09] Rolf Krause. A nonsmooth multiscale method for solving frictional two-body contact problems in 2d and 3d with multigrid efficiency. *SIAM Journal on Scientific Computing*, 31(2):1399–1423, 2009.

[KS80] David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 31. Siam, 1980.

[KS17] Bernhard Kober and Gerhard Starke. Strong vs. weak symmetry in stress-based mixed finite element methods for linear elasticity. In *European Conference on Numerical Mathematics and Advanced Applications*, pages 323–333. Springer, 2017.

[KVW15] Rolf Krause, Andreas Veeser, and Mirjam Walloth. An efficient and reliable residual-type a posteriori error estimator for the signorini problem. *Numerische Mathematik*, 130(1):151–197, 2015.

[LM72] J. L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications, I.* Springer-Verlag, Berlin-Heidelberg-New York, 1972.

[LS67] Jacques-Louis Lions and Guido Stampacchia. Variational inequalities. *Communications on pure and applied mathematics*, 20(3):493–519, 1967.

[LV04] Marco Lonsing and Rüdiger Verfürth. A posteriori error estimators for mixed finite element methods in linear elasticity. *Numerische Mathematik*, 97(4):757–778, 2004.

[Neč12] J. Nečas. *Direct Methods in the Theory of Elliptic Equations*. Springer, Berlin, 2012.

[NW06] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[OK80] John Tinsley Oden and Noboru Kikuchi. Theory of variational inequalities with applications to problems of flow through porous media. *International Journal of Engineering Science*, 18(10):1173–1284, 1980.

[PS47] William Prager and John L Synge. Approximations in elasticity based on the concept of function space. *Quarterly of Applied Mathematics*, 5(3):241–269, 1947.

[Ren06] Yves Renard. A uniqueness criterion for the signorini problem with coulomb friction. *SIAM journal on mathematical analysis*, 38(2):452–467, 2006.

[RKSK18] Gabriele Rovi, Bernhard Kober, Gerhard Starke, and Rolf Krause. Monotone multilevel for fosls linear elastic contact. *Technical report, ICS Report*, 2018.

[Rös00] Andreas Rössle. Corner singularities and regularity of weak solutions for the two-dimensional lamé equations on domains with angular corners. *Journal of elasticity and the physical science of solids*, 60(1):57–75, 2000.

[Sig33] Antonio Signorini. Sopra alcune questioni di statica dei sistemi continui. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 2(2):231–251, 1933.

[Sta07] Gerhard Starke. An adaptive least-squares mixed finite element method for elasto-plasticity. *SIAM Journal on Numerical Analysis*, 45(1):371–388, 2007.

[Ste91] R. Stenberg. Postprocessing schemes for some mixed finite elements. *Math. Model. Numer. Anal.*, 25:151–167, 1991.

[Voh10] M. Vohralík. Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods. *Math. Comp.*, 79:2001–2032, 2010.