

GENERALIZED NEWTON METHODS FOR THE 2D-SIGNORINI CONTACT PROBLEM WITH FRICTION IN FUNCTION SPACE

KARL KUNISCH¹ AND GEORG STADLER¹

Abstract. The 2D-Signorini contact problem with Tresca and Coulomb friction is discussed in infinite-dimensional Hilbert spaces. First, the problem with given friction (Tresca friction) is considered. It leads to a constraint non-differentiable minimization problem. By means of the Fenchel duality theorem this problem can be transformed into a constrained minimization involving a smooth functional. A regularization technique for the dual problem motivated by augmented Lagrangians allows to apply an infinite-dimensional semi-smooth Newton method for the solution of the problem with given friction. The resulting algorithm is locally superlinearly convergent and can be interpreted as active set strategy. Combining the method with an augmented Lagrangian method leads to convergence of the iterates to the solution of the original problem. Comprehensive numerical tests discuss, among others, the dependence of the algorithm's performance on material and regularization parameters and on the mesh. The remarkable efficiency of the method carries over to the Signorini problem with Coulomb friction by means of fixed point ideas.

Mathematics Subject Classification. 49M05, 49M29, 74M10, 74M15, 74B05.

Received: July 13, 2004. Revised: March 4, 2005.

1. INTRODUCTION

This paper is concerned with the development and convergence analysis of algorithms for the solution of 2D-contact problems with Tresca and Coulomb friction. The main difficulty of these problems lies in the contact and friction conditions, which are inherently nonlinear and non-differentiable thus making both theoretical analysis as well as efficient numerical realization truly challenging.

In contact problems, also known as Signorini problems, one has to detect the contact zone between an elastic body and a rigid foundation that is *a priori* unknown. At the contact boundary, frictional forces are often too large to be neglected. Thus, besides the non-penetration condition, one also has to take into account the frictional behavior in the contact zone. The predominant friction laws used in literature are the Tresca and Coulomb law. While the contact problem with Tresca friction leads to a classical variational inequality, the Coulomb friction problem results in a quasi-variational inequality. This makes proving theoretical results for Coulomb's friction difficult or even impossible, *e.g.*, a solution to the quasistatic contact problem with Coulomb friction exists only if the friction coefficient is sufficiently small, [9, 15].

Keywords and phrases. Signorini contact problems, Coulomb and Tresca friction, linear elasticity, semi-smooth Newton method, Fenchel dual, augmented Lagrangians, complementarity system, active sets.

¹ Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria. karl.kunisch@uni-graz.at; ge.stadler@uni-graz.at

© EDP Sciences, SMAI 2005

Instead of the Coulomb law, frequently the simpler Tresca friction law is used. Moreover, a commonly chosen approach towards the solution of Coulomb frictional contact problems is to define the solution as a fixed point of a sequence of solutions to the Tresca problem.

Concerning the numerical realization of Coulomb frictional contact problems, two main approaches can be found in the literature. First, one may treat the discretized system directly; we refer to [1, 6, 7, 25, 28] for Newton-type methods. A drawback of this approach is that it only applies in finite dimensions and that convergence results are difficult to obtain. However, the authors of the above mentioned papers report on good numerical results. The second and more commonly used approach is to utilize a sequence of Tresca friction problems combined with a fixed point concept (see, *e.g.*, [16, 26, 29, 32]). Thus, the crucial requirement to obtain an efficient numerical algorithm for Coulomb frictional contact problems lies in a fast and reliable algorithm for the solution of the Signorini problem with Tresca friction. Besides the fact that frequently the Tresca friction model itself is used in applications, this motivates the development of fast solvers for contact problems with Tresca friction.

Early papers concerned with this question are [29, 31]. They utilize a successive overrelaxation method and the Uzawa algorithm to solve the Tresca friction problem, see also the review article [32]. The authors of the more recent articles [8, 16] use a discrete dual formulation of the problem and quadratic programming methods with proportioning and projections for the solution of discrete 2D-Tresca frictional contact problems. A different idea is followed in [26], where monotone multigrid methods are applied to construct an efficient and globally convergent solver for discrete Tresca frictional contact problems. The implementation of this method is rather complicated and only few convergence rate results are available.

The present paper is devoted to the development and analysis of algorithms for 2D-frictional contact problems in *infinite* dimensional function spaces. A recent generalized differentiability concept in a Hilbert space framework (see the definitions below) is applied to derive second order methods for elasticity problems subject to unilateral contact with friction. The approach taken here is to a large extent based on the Fenchel duality theorem (that we also summarize below) that allows to transform a non-differentiable minimization problem into an inequality constrained minimization of a smooth functional. This approach is applied to the contact problem with Tresca friction, which can be formulated as constraint minimization of a non-differentiable functional. Aside from using just the first order necessary conditions of this problem, which are usually the starting points of the analysis, we additionally use for our investigation alternately the primal and dual formulations of the problem. Another important aspect in this paper is the use of certain nonlinear complementarity functions that allow to write complementarity conditions as nonsmooth operator equations in function spaces. An application of the semi-smooth Newton methods as developed in [18, 23, 35] to the (smoothed) set of necessary optimality conditions leads to new, superlinearly convergent algorithms for the solution of 2D-contact problems with Tresca friction. The methods turn out to be related to the primal-dual active set strategy as used for optimal control problems (see, *e.g.*, [4, 18]), and for discrete unilaterally constrained variational problems and pure contact problems, [19, 21]. The regularization that is used is motivated by augmented Lagrangians (see [1, 22, 25]). However, in this paper we mainly aim at the generalized Newton method for the solution of the regularized problem with (possibly) large regularization parameters rather than at (first-order) augmented Lagrangian algorithms, which are briefly discussed in Section 5 for reasons of completeness.

We now give a formulation of the Signorini contact problem with Coulomb friction in 2D. For this purpose we consider an elastic body that occupies in its initial configuration the open and bounded domain $\Omega \subset \mathbb{R}^2$ with $\mathcal{C}^{1,1}$ -boundary $\Gamma = \partial\Omega$. Let this boundary be divided into three disjoint parts, namely the Dirichlet part Γ_d , further the part Γ_n with prescribed surface loads $\mathbf{h} \in \mathbf{L}^2(\Gamma_n) := (L^2(\Gamma_n))^2$ and the part Γ_c , where contact and friction with a rigid foundation may occur. For simplicity we assume that $\bar{\Gamma}_c \cap \bar{\Gamma}_d = \emptyset$ to avoid working with the space $H_{00}^{1/2}(\Gamma_c)$. We are interested in the deformation $\mathbf{y} = (y_1, y_2)^\top$ of the elastic body which is also subject to a given body force $\mathbf{f} \in \mathbf{L}^2(\Omega) := (L^2(\Omega))^2$. The gap between elastic body and rigid foundation is $d := \tau_N \mathbf{d} \geq 0$, where $\mathbf{d} \in \mathbf{H}^1(\Omega) := (H^1(\Omega))^2$ and $\tau_N \mathbf{y}$ denotes the normal component of the trace along Γ_c . Denoting the linear strain and stress tensors for \mathbf{y} by $\boldsymbol{\varepsilon} \mathbf{y}$ and $\boldsymbol{\sigma} \mathbf{y}$ as common in plane elasticity (for their definition see Sect. 2),

the Signorini problem with Coulomb friction is given as follows:

$$-\text{Div } \boldsymbol{\sigma}(\mathbf{y}) = \mathbf{f} \quad \text{in } \Omega, \tag{1.1a}$$

$$\boldsymbol{\tau} \mathbf{y} = 0 \quad \text{on } \Gamma_d, \tag{1.1b}$$

$$(\boldsymbol{\sigma} \mathbf{y}) \mathbf{n} = \mathbf{h} \quad \text{on } \Gamma_n, \tag{1.1c}$$

$$\tau_N \mathbf{y} - d \leq 0, \quad \sigma_N \mathbf{y} \leq 0, \quad (\tau_N \mathbf{y} - d) \sigma_N \mathbf{y} = 0 \quad \text{on } \Gamma_c, \tag{1.1d}$$

$$|\sigma_T \mathbf{y}| < \mathfrak{F} |\sigma_N \mathbf{y}| \quad \text{on } \{x \in \Gamma_c : \tau_T \mathbf{y} = 0\}, \tag{1.1e}$$

$$\sigma_T \mathbf{y} = -\mathfrak{F} \frac{|\sigma_N \mathbf{y}|}{|\tau_T \mathbf{y}|} \tau_T \mathbf{y} \quad \text{on } \{x \in \Gamma_c : \tau_T \mathbf{y} \neq 0\}. \tag{1.1f}$$

Above, “Div” denotes the row-wise divergence-operator and we denote by $\boldsymbol{\tau} : \mathbf{H}^1(\Omega) \rightarrow \mathbf{H}^{1/2}(\Gamma) := (H^{1/2}(\Gamma))^2$ the usual zero-order trace mapping. The corresponding scalar valued normal and tangential component mappings are denoted by $\tau_N, \tau_T : \mathbf{H}^1(\Omega) \rightarrow H^{1/2}(\Gamma_c)$, *i.e.*, for $\mathbf{y} \in \mathbf{H}^1(\Omega)$ we have the splitting $\boldsymbol{\tau} \mathbf{y} = (\tau_N \mathbf{y}) \mathbf{n} + (\tau_T \mathbf{y}) \mathbf{t}$ with \mathbf{n} and \mathbf{t} denoting the unit normal and tangential vector along Γ_c , respectively. Similarly, using (1.1a) we can, following [24], decompose the stress along the boundary, namely $(\boldsymbol{\sigma} \mathbf{y}) \mathbf{n} = (\sigma_N \mathbf{y}) \mathbf{n} + (\sigma_T \mathbf{y}) \mathbf{t}$ with mappings $\sigma_N, \sigma_T : \mathbf{Y} \rightarrow H^{-1/2}(\Gamma_c)$. Moreover, $\mathfrak{F} : \Gamma_c \rightarrow \mathbb{R}$ denotes the friction coefficient, whose regularity is discussed in the next section.

There are major mathematical difficulties inherent in the problem (1.1). For instance, (1.1) cannot be associated to an optimization problem for which standard *a priori* estimates would guarantee existence or uniqueness of a solution. Further, in general $\sigma_N \mathbf{y}$ in (1.1e), (1.1f) is not pointwise almost everywhere defined which makes the definition of $|\sigma_N \mathbf{y}|$ and thus a mathematical precise formulation of (1.1) difficult. Replacing the Coulomb friction in the above model by Tresca friction means replacing $|\sigma_N \mathbf{y}|$ by a given friction g . Doing so, the resulting system can be analyzed and the existence of a unique solution can be proved. Furthermore, a reformulation of the problem can be derived using the Fenchel duality theorem in Banach spaces (see, *e.g.*, [10]) that we now recall for the reader’s convenience.

For this purpose let V and Y be Banach spaces with topological duals V^* and Y^* , respectively. Further let $\Lambda \in \mathcal{L}(V, Y)$, the space of continuous linear operators from V to Y , and let $\mathcal{F} : V \rightarrow \mathbb{R} \cup \{\infty\}$, $\mathcal{G} : Y \rightarrow \mathbb{R} \cup \{\infty\}$ be convex, proper and lower semicontinuous such that there exists $v_0 \in V$ with $\mathcal{F}(v_0) < \infty$, $\mathcal{G}(\Lambda v_0) < \infty$ and \mathcal{G} is continuous at Λv_0 . Then,

$$\inf_{v \in V} \{\mathcal{F}(v) + \mathcal{G}(\Lambda v)\} = \sup_{q \in Y^*} \{-\mathcal{F}^*(-\Lambda^* q) - \mathcal{G}^*(q)\}, \tag{1.2}$$

where $\Lambda^* \in \mathcal{L}(Y^*, V^*)$ is the adjoint of Λ . The convex conjugates $\mathcal{F}^* : V^* \rightarrow \mathbb{R} \cup \{\infty\}$, $\mathcal{G}^* : Y^* \rightarrow \mathbb{R} \cup \{\infty\}$ of \mathcal{F} and \mathcal{G} , respectively, are defined by

$$\mathcal{F}^*(v^*) = \sup_{v \in V} \{\langle v, v^* \rangle_{V, V^*} - \mathcal{F}(v)\}, \tag{1.3}$$

and analogously for \mathcal{G}^* . The conditions imposed on \mathcal{F} and \mathcal{G} guarantee that the dual problem (*i.e.*, the problem on the right hand side of (1.2)) admits a solution. Furthermore, the solutions $\bar{v} \in V$ and $\bar{q} \in Y^*$ are characterized by the extremality conditions

$$\begin{aligned} -\Lambda^* \bar{q} &\in \partial \mathcal{F}(\bar{v}), \\ \bar{q} &\in \partial \mathcal{G}(\Lambda \bar{v}), \end{aligned} \tag{1.4}$$

where ∂ denotes the subdifferential from convex analysis.

Reformulating the Tresca friction problems by means of the Fenchel duality theorem will motivate the application of a semi-smooth Newton method to the resulting system. We conclude this introduction by summarizing

the main definitions and results for this infinite-dimensional semi-smooth Newton approach, where we follow [18] (see also [5, 35]).

Let X, Y be Banach spaces and $F : D \subset X \rightarrow Y$ be a nonlinear mapping with open domain D . Then, we call the mapping F *Newton differentiable on the open subset* $U \subset D$ if there exists a mapping $G : U \rightarrow \mathcal{L}(X, Y)$ such that

$$\lim_{h \rightarrow 0} \frac{1}{\|h\|} \|F(x+h) - F(x) - G(x+h)h\| = 0$$

for every $x \in U$. The mapping G in the above definition is referred to as *generalized derivative*. Note that G is not necessarily unique. In [18] it is shown that the pointwise max- and min-operators are Newton differentiable if considered as mapping between appropriate function spaces. To be precise, let X denote a function space of real-valued functions on some $\Omega \subset \mathbb{R}^n$, further $\max(0, y)$ and $\min(0, y)$ the pointwise max- and min-operations, respectively. Defining the following candidates for the generalized derivatives

$$G_{\max}(y)(x) = \begin{cases} 1 & \text{if } y(x) \geq 0, \\ 0 & \text{if } y(x) < 0; \end{cases} \quad G_{\min}(y)(x) = \begin{cases} 1 & \text{if } y(x) \leq 0, \\ 0 & \text{if } y(x) > 0, \end{cases} \quad (1.5)$$

one can prove the following result.

Theorem 1.1. *The mappings $\max(0, \cdot) : L^q(\Omega) \rightarrow L^p(\Omega)$ and $\min(0, \cdot) : L^q(\Omega) \rightarrow L^p(\Omega)$ with $1 \leq p < q < \infty$ are Newton differentiable on $L^q(\Omega)$ with generalized derivatives G_{\max} and G_{\min} , respectively.*

Note that Theorem 1.1 requires a norm gap (*i.e.*, $p < q$) to hold true. In [18] it is shown that the functions in (1.5) cannot serve as generalized derivatives if $p \geq q$. For the Newton-differentiable equation $F(x) = 0$ we now consider the Newton step

$$x^{k+1} = x^k - G(x^k)^{-1}F(x^k), \quad (1.6)$$

where G is a generalized derivative in the sense of the above definition. Then, following [5, 18, 35] we have the following local convergence result that will be used later.

Theorem 1.2. *Suppose that $\bar{x} \in D$ is a solution to $F(x) = 0$ and that F is Newton differentiable in an open neighborhood U of \bar{x} and that $\{\|G(x)^{-1}\| : x \in U\}$ is bounded. Then the Newton-iteration (1.6) converges superlinearly to \bar{x} provided that $\|x^0 - \bar{x}\|$ is sufficiently small.*

Let us now briefly outline the structure of this paper. Following this introduction, the Signorini problem with Tresca friction is analyzed, its Fenchel dual and the corresponding extremality conditions are derived. In Section 3, we investigate a regularization procedure for the dual problem, derive the corresponding primal problem and the extremality conditions, and prove convergence of the solutions as the regularization parameters tend to infinity. In Sections 4 and 5, a semi-smooth Newton and an augmented Lagrangian method for the solution of the regularized problem are presented and analyzed in a Hilbert space framework and the relation to active set strategies is discussed. In Section 6, existence of a solution to the smoothed Signorini problem with Coulomb friction is proved and two fixed point techniques for the numerical realization of this problem are presented. Section 7 contains our numerical tests of the algorithms for Tresca as well as Coulomb frictional contact problems. In the concluding Section 8 we draw conclusions and give an outlook on further work.

2. CONTACT PROBLEM WITH GIVEN FRICTION

Due to the above mentioned problems with formulating and analyzing the contact problem with Coulomb friction, often the contact problem with given friction, also known as Tresca friction problem is considered (see, *e.g.*, [14, 20]). In this friction law the bound between slip and stick is given *a priori* and does not depend on \mathbf{y} such as in (1.1e), (1.1f). Then, the problem can be stated as optimization problem, which allows the application of arguments from convex analysis to argue existence of a unique solution. Moreover, the Tresca problem can also be utilized to obtain a mathematical precise weak formulation for the contact problem with Coulomb friction. This formulation relies on a fixed point argument that can also be exploited numerically to calculate solutions for the model with Coulomb friction.

2.1. Problem statement

We now give a precise statement of the Signorini problem with given friction. In the sequel we denote by (\cdot, \cdot) the L^2 -products and by $\langle \cdot, \cdot \rangle$ duality products. The set of admissible deformations is defined as

$$\mathbf{Y} := \{ \mathbf{v} \in \mathbf{H}^1(\Omega) : \boldsymbol{\tau} \mathbf{v} = 0 \text{ a.e. on } \Gamma_d \},$$

where here and below $\mathbf{H}^1(\Omega) := (H^1(\Omega))^2$. To incorporate the non-penetration condition between elastic body and rigid foundation we define the cone

$$\mathbf{K} := \{ \mathbf{v} \in \mathbf{Y} : \tau_N \mathbf{v} \leq 0 \text{ a.e. on } \Gamma_c \}.$$

As usual in linear elasticity, the linearized strain tensor is $\boldsymbol{\varepsilon}(\mathbf{y}) = \frac{1}{2} (\nabla \mathbf{y} + (\nabla \mathbf{y})^\top)$. Using Hooke’s law for the stress-strain relation, the linearized stress tensor $\boldsymbol{\sigma}(\mathbf{y}) := \mathbb{C} \boldsymbol{\varepsilon}(\mathbf{y}) := \lambda \text{tr}(\boldsymbol{\varepsilon}(\mathbf{y})) \text{Id} + 2\mu \boldsymbol{\varepsilon}(\mathbf{y})$ is obtained, where λ and μ are the Lamé parameters. These parameters are given by $\lambda = (E\nu)/((1+\nu)(1-2\nu))$ and $\mu = E/(2(1+\nu))$ with Young’s modulus $E > 0$ and the Poisson ration $\nu \in (0, 0.5)$. Above \mathbb{C} denotes the fourth-order isotropic material tensor for linear elasticity. We now define the symmetric bilinear form $a(\cdot, \cdot)$ on $\mathbf{Y} \times \mathbf{Y}$ and the linear form $L(\cdot)$ on \mathbf{Y} by

$$a(\mathbf{y}, \mathbf{z}) := \int_{\Omega} (\boldsymbol{\sigma} \mathbf{y}) : (\boldsymbol{\varepsilon} \mathbf{z}) \, dx, \quad L(\mathbf{y}) = \int_{\Omega} \mathbf{f} \mathbf{y} \, dx + \int_{\Gamma_n} \mathbf{h} \boldsymbol{\tau} \mathbf{y} \, dx,$$

where “:” denotes the sum of the componentwise products.

For given friction g we assume the regularity $g \in H^{-1/2}(\Gamma_c)$ and $g \geq 0$, i.e., $\langle g, h \rangle_{\Gamma_c} \geq 0$ for all $h \in H^{1/2}(\Gamma_c)$ with $h \geq 0$. Later in this section we will assume that g is more regular, namely $g \in L^2(\Gamma_c)$. However, for the following discussion we only require that $g \in H^{-1/2}(\Gamma_c)$, since, for the friction coefficient $\mathfrak{F} : \Gamma_c \rightarrow \mathbb{R}$ we assume that $\mathfrak{F} \in L^\infty(\Gamma_c)$ and furthermore that \mathfrak{F} belongs to the space of factors on $H^{1/2}(\Gamma_c)$, i.e., the mapping

$$H^{\frac{1}{2}}(\Gamma_c) \ni \lambda \mapsto \mathfrak{F} \lambda \in H^{\frac{1}{2}}(\Gamma_c)$$

is well-defined and bounded. By duality it follows that \mathfrak{F} is a factor on $H^{-1/2}(\Gamma_c)$ as well. In [13], p. 21, it is shown that, if \mathfrak{F} is uniformly Lipschitz continuous, it is a factor on $H^{1/2}(\Gamma_c)$. Finally, we define the following non-differentiable functional:

$$j(\mathbf{y}) := \int_{\Gamma_c} \mathfrak{F} g |\tau_T \mathbf{y}| \, dx.$$

After these preparations we can state the contact problem with given friction as minimization of a non-smooth functional over the set of admissible deformations, i.e., as

$$\min_{\mathbf{y} \in \mathbf{d} + \mathbf{K}} J(\mathbf{y}) := \frac{1}{2} a(\mathbf{y}, \mathbf{y}) - L(\mathbf{y}) + j(\mathbf{y}), \tag{P}$$

or equivalently as elliptic variational inequality [12]:

$$\begin{cases} \text{Find } \mathbf{y} \in \mathbf{d} + \mathbf{K} \text{ such that} \\ a(\mathbf{y}, \mathbf{z} - \mathbf{y}) + j(\mathbf{z}) - j(\mathbf{y}) \geq L(\mathbf{z} - \mathbf{y}) \text{ for all } \mathbf{z} \in \mathbf{d} + \mathbf{K}. \end{cases} \tag{2.1}$$

Due to the Korn inequality, the functional $J(\cdot)$ is uniformly convex, further it is lower semicontinuous. This implies that (P) and equivalently problem (2.1) admit a unique solution $\bar{\mathbf{y}} \in \mathbf{d} + \mathbf{K}$.

2.2. Dual problem

To derive the dual problem corresponding to (\mathcal{P}) , we apply the Fenchel calculus to the mappings $\mathcal{F} : \mathbf{Y} \rightarrow \mathbb{R}$ and $\mathcal{G} : \mathbf{V} \times H^{1/2}(\Gamma_c) \rightarrow \mathbb{R}$ given by

$$\mathcal{F}(\mathbf{y}) := \begin{cases} -L(\mathbf{y}) & \text{if } \mathbf{y} \in \mathbf{d} + \mathbf{K}, \\ \infty & \text{else;} \end{cases} \quad \mathcal{G}(\mathbf{q}, \nu) := \frac{1}{2} \int_{\Omega} \mathbf{q} : \mathbb{C} \mathbf{q} \, dx + \int_{\Gamma_c} \mathfrak{F}g|\nu| \, dx,$$

where

$$\mathbf{V} = \{\mathbf{p} \in (L^2(\Omega))^{2 \times 2} : p_{12} = p_{21}\}.$$

Furthermore, $\Lambda \in \mathcal{L}(\mathbf{Y}, \mathbf{V} \times H^{1/2}(\Gamma_c))$ is given by

$$\Lambda \mathbf{y} := (\Lambda_1 \mathbf{y}, \Lambda_2 \mathbf{y}) = (\varepsilon \mathbf{y}, \tau_T \mathbf{y}),$$

which allows to express (\mathcal{P}) as

$$\min_{\mathbf{y} \in \mathbf{Y}} \{\mathcal{F}(\mathbf{y}) + \mathcal{G}(\Lambda \mathbf{y})\}.$$

Endowing $\mathbf{V} \times H^{1/2}(\Gamma_c)$ with the usual product norm, it is easy to see that \mathcal{F} and \mathcal{G} satisfy the conditions for the Fenchel duality theorem, yielding that (1.2) and the extremality conditions (1.4) hold and characterize the solutions of primal and dual problem. We now calculate the convex conjugate functions $\mathcal{F}^*, \mathcal{G}^*$ corresponding to \mathcal{F}, \mathcal{G} explicitly. Following the definitions of the convex conjugate one derives that $\mathcal{F}^*(-\Lambda^*(\mathbf{p}, \mu))$ equals $+\infty$ unless

$$-\text{Div } \mathbf{p} = \mathbf{f}, \quad \mathbf{p} \cdot \mathbf{n} = \mathbf{h} \text{ in } \mathbf{L}^2(\Gamma_n) \quad \text{and} \quad p_T + \mu = 0 \text{ in } H^{-\frac{1}{2}}(\Gamma_c), \tag{2.2}$$

where $p_T = (\mathbf{n}^\top \mathbf{p}) \cdot \mathbf{t} \in H^{-1/2}(\Gamma_c)$. Further, one derives that

$$\mathcal{F}^*(-\Lambda^*(\mathbf{p}, \mu)) = \begin{cases} -\langle p_N, d \rangle_{\Gamma_c} & \text{if (2.2) and } p_N \leq 0 \text{ in } H^{-\frac{1}{2}}(\Gamma_c) \text{ hold,} \\ \infty & \text{else.} \end{cases}$$

Evaluating the convex conjugate for \mathcal{G} yields that

$$\mathcal{G}^*(\mathbf{p}, \mu) = \begin{cases} \frac{1}{2} \int_{\Omega} \mathbb{C}^{-1} \mathbf{p} : \mathbf{p} \, dx & \text{if } \langle \mathfrak{F}g, |\nu| \rangle_{\Gamma_c} - \langle \nu, \mu \rangle_{\Gamma_c} \geq 0 \text{ for all } \nu \in H^{\frac{1}{2}}(\Gamma_c), \\ \infty & \text{else.} \end{cases}$$

Thus, following (1.2) we derive the dual problem corresponding to (\mathcal{P}) :

$$\begin{aligned} & \sup && -\frac{1}{2} \int_{\Omega} \mathbb{C}^{-1} \mathbf{p} : \mathbf{p} \, dx + \langle p_N, d \rangle_{\Gamma_c}. && (\mathcal{P}^*) \\ & (\mathbf{p}, \mu) \in \mathbf{V} \times H^{-\frac{1}{2}}(\Gamma_c) && && \\ & \text{s.t. (2.2), } p_N \leq 0 \text{ in } H^{-\frac{1}{2}}(\Gamma_c), && && \\ & \text{and } \langle \mathfrak{F}g, |\nu| \rangle_{\Gamma_c} - \langle \nu, \mu \rangle_{\Gamma_c} \geq 0 && && \\ & \text{for all } \nu \in H^{\frac{1}{2}}(\Gamma_c). && && \end{aligned}$$

This problem is a constrained maximization problem of a quadratic functional, while the primal problem (\mathcal{P}) involves the minimization of a non-differentiable functional. Evaluating the extremality conditions (1.4) for the above problems one obtains the following lemma.

Lemma 2.1. *The solution $\bar{\mathbf{y}} \in \mathbf{d} + \mathbf{K}$ of (\mathcal{P}) and the solution $(\bar{\mathbf{p}}, \bar{\mu})$ of (\mathcal{P}^*) are characterized by $\boldsymbol{\sigma}\bar{\mathbf{y}} = \bar{\mathbf{p}}$ and the existence of $\bar{\lambda} \in H^{-1/2}(\Gamma_c)$ such that*

$$a(\bar{\mathbf{y}}, \mathbf{z}) - L(\mathbf{z}) + \langle \bar{\mu}, \tau_T \mathbf{z} \rangle_{\Gamma_c} + \langle \bar{\lambda}, \tau_N \mathbf{z} \rangle_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}, \tag{2.3a}$$

$$\langle \bar{\lambda}, \tau_N \mathbf{z} \rangle_{\Gamma_c} \leq 0 \text{ for all } \mathbf{z} \in \mathbf{K}, \tag{2.3b}$$

$$\langle \bar{\lambda}, \tau_N \bar{\mathbf{y}} - d \rangle_{\Gamma_c} = 0, \tag{2.3c}$$

$$\langle \bar{\mathfrak{F}}g, |\nu| \rangle_{\Gamma_c} - \langle \bar{\mu}, \nu \rangle_{\Gamma_c} \geq 0 \text{ for all } \nu \in H^{\frac{1}{2}}(\Gamma_c), \tag{2.3d}$$

$$\langle \bar{\mathfrak{F}}g, |\tau_T \bar{\mathbf{y}}| \rangle_{\Gamma_c} - \langle \bar{\mu}, \tau_T \bar{\mathbf{y}} \rangle_{\Gamma_c} = 0. \tag{2.3e}$$

Proof. Evaluating the extremality condition $-\Lambda^*(\bar{\mathbf{p}}, \bar{\mu}) \in \partial\mathcal{F}(\bar{\mathbf{y}})$ results in $\bar{\mathbf{y}} \in \mathbf{d} + \mathbf{K}$ and

$$(\bar{\mathbf{p}}, \boldsymbol{\varepsilon}(\mathbf{z} - \bar{\mathbf{y}})) - L(\mathbf{z} - \bar{\mathbf{y}}) + \langle \bar{\mu}, \tau_T(\mathbf{z} - \bar{\mathbf{y}}) \rangle_{\Gamma_c} \geq 0 \text{ for all } \mathbf{z} \in \mathbf{d} + \mathbf{K}. \tag{2.4}$$

Condition $(\bar{\mathbf{p}}, \bar{\mu}) \in \partial\mathcal{G}(\Lambda\bar{\mathbf{y}})$ yields that $\bar{\mathbf{p}} = \boldsymbol{\sigma}\bar{\mathbf{y}}$ and the equations (2.3d) and (2.3e). Introducing the multiplier $\bar{\lambda}$ for the variational inequality (2.4) leads to (2.3a), (2.3b) and (2.3c). \square

Note that, according to (2.2), for the multiplier $\bar{\mu}$ corresponding to the non-differentiability of the primal functional $J(\cdot)$ we have the mechanical interpretation $\bar{\mu} = -\sigma_T \bar{\mathbf{y}}$. Using Green's theorem in (2.3a), one also finds a mechanical interpretation for the above introduced multiplier $\bar{\lambda}$ corresponding to the contact condition, namely, $\bar{\lambda}$ is the negative stress in normal direction, *i.e.*,

$$\bar{\lambda} = -\sigma_N \bar{\mathbf{y}}. \tag{2.5}$$

We now briefly comment on the case that the given friction g is more regular, namely $g \in L^2(\Gamma_c)$. In this case we can define \mathcal{G} on the larger set $\mathbf{V} \times L^2(\Gamma_c)$. One can verify that the assumptions for the Fenchel duality theorem hold, and thus we obtain higher regularity for the dual variable μ corresponding to the non-differentiability of the cost functional in (\mathcal{P}^*) , namely $\mu \in L^2(\Gamma_c)$. This implies that the dual problem can be written as follows:

$$\begin{aligned} & \sup_{(\mathbf{p}, \mu) \in \mathbf{V} \times L^2(\Gamma_c)} -\frac{1}{2} \int_{\Omega} \mathbb{C}^{-1} \mathbf{p} : \mathbf{p} \, dx + \langle p_N, d \rangle_{\Gamma_c}. \tag{2.6} \\ & \text{s.t. (2.2), } p_N \leq 0 \text{ in } H^{-\frac{1}{2}}(\Gamma_c), \\ & \text{and } |\mu| \leq \bar{\mathfrak{F}}g \text{ a.e. on } \Gamma_c. \end{aligned}$$

Utilizing the relation between primal and dual variables, in particular $\mathbf{p} = \boldsymbol{\sigma}\mathbf{y}$ and (2.5), one can transform (2.6) into

$$\left\{ \begin{array}{l} - \min_{\substack{\lambda \geq 0 \text{ in } H^{-\frac{1}{2}}(\Gamma_c) \\ |\mu| \leq \bar{\mathfrak{F}}g \text{ a.e. on } \Gamma_c}} \frac{1}{2} a(\mathbf{y}_{\lambda, \mu}, \mathbf{y}_{\lambda, \mu}) + \langle \lambda, d \rangle_{\Gamma_c}, \\ \text{where } \mathbf{y}_{\lambda, \mu} \text{ satisfies} \\ a(\mathbf{y}_{\lambda, \mu}, \mathbf{z}) - L(\mathbf{z}) + \langle \lambda, \tau_N \mathbf{z} \rangle_{\Gamma_c} + \langle \mu, \tau_T \mathbf{z} \rangle_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}. \end{array} \right. \tag{2.7}$$

Note that problem (2.7) is an equivalent form for the dual problem (2.6), now written in the variables λ and μ . The primal variable $\mathbf{y}_{\lambda, \mu}$ appears only as auxiliary variable determined from λ and μ . Since $g \in L^2(\Gamma_c)$, also the extremality conditions corresponding to (\mathcal{P}) and (2.6) can be given more explicitly. First, (2.3d) is equivalent to

$$|\bar{\mu}| \leq \bar{\mathfrak{F}}g \text{ a.e. on } \Gamma_c, \tag{2.3d'}$$

and a brief computation shows that (2.3e) is equivalent to

$$\begin{cases} \tau_T \bar{\mathbf{y}} = 0 \text{ or} \\ \tau_T \bar{\mathbf{y}} \neq 0 \text{ and } \bar{\mu} = \mathfrak{F}g \frac{\tau_T \bar{\mathbf{y}}}{|\tau_T \bar{\mathbf{y}}|}. \end{cases} \tag{2.3e'}$$

By a short calculation one finds that (2.3d') and (2.3e') can equivalently be expressed as

$$\sigma \tau_T \bar{\mathbf{y}} - \max(0, \sigma \tau_T \bar{\mathbf{y}} + \bar{\mu} - \mathfrak{F}g) - \min(0, \sigma \tau_T \bar{\mathbf{y}} + \bar{\mu} + \mathfrak{F}g) = 0, \tag{2.8}$$

with arbitrary $\sigma > 0$.

2.3. Weak formulation of the Coulomb friction problem

Having the above results available we can now give a weak formulation of the contact problem with Coulomb friction that circumvents the problems mentioned in the introduction, namely the lack of regularity of $\sigma_N \mathbf{y}$ in (1.1). The formulation given below utilizes the contact problem with given friction $g \in H^{-1/2}(\Gamma_c)$ and a fixed point idea. We define the cone of nonnegative functionals over $H^{1/2}(\Gamma_c)$ as

$$H_+^{-\frac{1}{2}}(\Gamma_c) := \{ \xi \in H^{-\frac{1}{2}}(\Gamma_c) : \langle \xi, \eta \rangle_{\Gamma_c} \geq 0 \text{ for all } \eta \in H^{\frac{1}{2}}(\Gamma_c), \eta \geq 0 \}.$$

Then, we consider the mapping $\Psi : H_+^{-1/2}(\Gamma_c) \rightarrow H_+^{-1/2}(\Gamma_c)$ defined by $\Psi(g) := \lambda_g$, where λ_g is the unique multiplier for the contact condition in (2.3) for the problem with given friction g . Property (2.3b) implies that Ψ is well-defined. This allows us, having (2.5) in mind, to call $\mathbf{y} \in \mathbf{Y}$ weak solution of the Signorini problem with Coulomb friction if its negative normal boundary stress $-\sigma_N \mathbf{y}$ is a fixed point of the mapping Ψ . In general, such a fixed point for the mapping Ψ does not exist, *i.e.*, the Coulomb friction problem does not always have a solution. It is proved in [9, 15, 20, 30] that the contact problem with Coulomb friction admits a weak solution if the friction coefficient \mathfrak{F} is sufficiently small.

3. THE REGULARIZED CONTACT PROBLEM WITH TRESCA FRICTION

We now introduce and analyze a regularized version of the contact problem with given friction that allows the application of the generalized Newton method from Section 1. For that purpose we intend to write the complementarity conditions (2.3b), (2.3c) and (2.3d), (2.3e) as nonlinear equations using complementarity functions. This is only possible if $\lambda, \bar{\mu} \in L^2(\Gamma_c)$ and therefore requires proper regularization. Another reason for the necessity of introducing a regularization can be seen from considering (2.8). Following Theorem 1.2, for Newton differentiability of the max- and min-function a norm norm gap is required. For the variable $\bar{\mu}$ that appears under the max- and min-functions we cannot expect any smoothing leading to the required norm gap. Using the regularization technique below, which is motivated by augmented Lagrangians, the explicit appearance of $\bar{\mu}$ in the max- and min- function can be avoided.

In the sequel we assume $g \in L^2(\Gamma_c)$. Motivated from the above discussion and results in [34], we start our consideration with a regularized version of the dual problem (2.6) written in the form (2.7). For this purpose, for $\gamma_1, \gamma_2 > 0$, given $\hat{\lambda} \in L^2(\Gamma_c)$ and $\hat{\mu} \in L^2(\Gamma_c)$ we define the functional $J_{\gamma_1, \gamma_2}^* : L^2(\Gamma_c) \times L^2(\Gamma_c) \rightarrow \mathbb{R}$ by

$$J_{\gamma_1, \gamma_2}^*(\lambda, \mu) := \frac{1}{2} a(\mathbf{y}_{\lambda, \mu}, \mathbf{y}_{\lambda, \mu}) + (\lambda, d)_{\Gamma_c} + \frac{1}{2\gamma_1} \|\lambda - \hat{\lambda}\|_{\Gamma_c}^2 + \frac{1}{2\gamma_2} \|\mu - \hat{\mu}\|_{\Gamma_c}^2 - \frac{1}{2\gamma_1} \|\hat{\lambda}\|_{\Gamma_c}^2 - \frac{1}{2\gamma_2} \|\hat{\mu}\|_{\Gamma_c}^2,$$

where $\mathbf{y}_{\lambda, \mu} \in \mathbf{Y}$ satisfies

$$a(\mathbf{y}_{\lambda, \mu}, \mathbf{z}) - L(\mathbf{z}) + (\lambda, \tau_N \mathbf{z})_{\Gamma_c} + (\mu, \tau_T \mathbf{z})_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}. \tag{3.1}$$

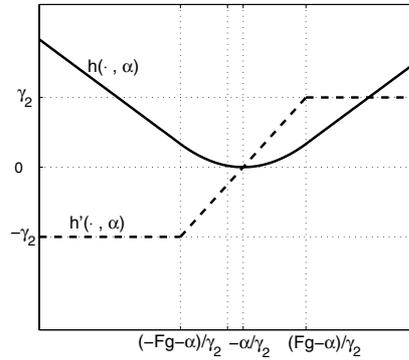


FIGURE 1. Illustration of the function $h(\cdot, \alpha)$ as defined in (3.2) and of its derivative $h'(\cdot, \alpha)$.

Now, the regularized dual problem with given friction is defined as

$$\max_{\lambda \geq 0, |\mu| \leq \mathfrak{F}g \text{ a.e. on } \Gamma_c} -J_{\gamma_1, \gamma_2}^*(\lambda, \mu). \tag{P_{\gamma_1, \gamma_2}^*}$$

Obviously, the last two terms in the definition of J_{γ_1, γ_2}^* are constants and can thus be neglected in the optimization problem $(P_{\gamma_1, \gamma_2}^*)$. However, they are introduced with regard to the primal problem corresponding to $(P_{\gamma_1, \gamma_2}^*)$, which we turn to next. We define the functional $J_{\gamma_1, \gamma_2} : \mathbf{Y} \rightarrow \mathbb{R}$ by

$$J_{\gamma_1, \gamma_2}(\mathbf{y}) := \frac{1}{2}a(\mathbf{y}, \mathbf{y}) - L(\mathbf{y}) + \frac{1}{2\gamma_1} \|\max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y} - d))\|_{\Gamma_c}^2 + \frac{1}{\gamma_2} \int_{\Gamma_c} \mathfrak{F}g h(\tau_T \mathbf{y}(x), \hat{\mu}(x)) \, dx,$$

where $h(\cdot, \cdot) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a local smoothing of the absolute value function, namely (see also Fig. 1)

$$h(x, \alpha) := \begin{cases} |\gamma_2 x + \alpha| - \frac{1}{2}\mathfrak{F}g & \text{if } |\gamma_2 x + \alpha| \geq \mathfrak{F}g, \\ \frac{1}{2\mathfrak{F}g} |\gamma_2 x + \alpha|^2 & \text{if } |\gamma_2 x + \alpha| < \mathfrak{F}g. \end{cases} \tag{3.2}$$

Then, the primal problem corresponding to $(P_{\gamma_1, \gamma_2}^*)$ is

$$\min_{\mathbf{y} \in \mathbf{Y}} J_{\gamma_1, \gamma_2}(\mathbf{y}). \tag{P_{\gamma_1, \gamma_2}}$$

This can be verified similarly as for the original problem using Fenchel duality theory (we refer to [33] for the detailed derivation). Clearly, both problems (P_{γ_1, γ_2}) and $(P_{\gamma_1, \gamma_2}^*)$ admit unique solutions $\mathbf{y}_{\gamma_1, \gamma_2}$ and $(\lambda_{\gamma_1, \gamma_2}, \mu_{\gamma_1, \gamma_2})$, respectively. Note that the regularization turns the primal problem into the unconstrained minimization of a continuously differentiable functional, while the corresponding dual problem is still the constrained minimization of a quadratic functional. To shorten notation we henceforth mark all variables of the regularized problems only by the index “ γ ” instead of “ γ_1, γ_2 ”. It can be shown that the extremality conditions relating (P_{γ_1, γ_2}) and $(P_{\gamma_1, \gamma_2}^*)$ are

$$a(\mathbf{y}_\gamma, \mathbf{z}) - L(\mathbf{z}) + (\mu_\gamma, \tau_T \mathbf{z})_{\Gamma_c} + (\lambda_\gamma, \tau_N \mathbf{z})_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}, \tag{3.3a}$$

$$\lambda_\gamma - \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_\gamma - d)) = 0 \text{ on } \Gamma_c, \tag{3.3b}$$

$$\begin{cases} \gamma_2(\xi_\gamma - \tau_T \mathbf{y}_\gamma) + \mu_\gamma - \hat{\mu} = 0, \\ \xi_\gamma - \max(0, \xi_\gamma + \sigma(\mu_\gamma - \mathfrak{F}g)) - \min(0, \xi_\gamma + \sigma(\mu_\gamma + \mathfrak{F}g)) = 0 \end{cases} \tag{3.3c}$$

for any $\sigma > 0$. Here, ξ_γ is the Lagrange multiplier associated to the constraint $|\mu| \leq \mathfrak{F}g$ in $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$. By setting $\sigma = \gamma_2^{-1}$, ξ_γ can be eliminated from (3.3c), which results in

$$\gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu} - \mu_\gamma - \max(0, \gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu} - \mathfrak{F}g) - \min(0, \gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu} + \mathfrak{F}g) = 0. \tag{3.4}$$

However, compared to (3.4) equation (3.3c) involves the additional variable σ . In [34] it is observed that considering (3.3c) rather than (3.4) motivates a slightly different algorithm that turns out to give better numerical results.

Next we investigate the convergence of the primal variable \mathbf{y}_γ as well as the dual variables $(\lambda_\gamma, \mu_\gamma)$ as the regularization parameters γ_1, γ_2 tend to infinity. For this purpose we denote by $\bar{\mathbf{y}}$ the solution of (\mathcal{P}) and by $(\bar{\lambda}, \bar{\mu})$ the solution to (\mathcal{P}^*) .

Theorem 3.1. *For all $\hat{\lambda} \in L^2(\Gamma_c)$, $\hat{\mu} \in L^2(\Gamma_c)$ and $g \in L^2(\Gamma_c)$, the primal variable \mathbf{y}_γ converges to $\bar{\mathbf{y}}$ strongly in \mathbf{Y} and the dual variables $(\lambda_\gamma, \mu_\gamma)$ converge to $(\bar{\lambda}, \bar{\mu})$ weakly in $H^{-1/2}(\Gamma_c) \times L^2(\Gamma_c)$ as $\gamma_1 \rightarrow \infty$ and $\gamma_2 \rightarrow \infty$.*

Proof. The proof of this theorem can be found in Appendix A. □

4. A SEMI-SMOOTH NEWTON METHOD FOR $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$

We now present the semi-smooth Newton method for the solution of the regularized problems $(\mathcal{P}_{\gamma_1, \gamma_2})$ and $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$. An explicit calculation of the formal Newton step for system (3.3) using (1.5) leads to an iterative strategy that allows interpretation as active set strategy. Related methods were described in [1, 6, 7, 25]. These Newton-type methods are all restricted to finite dimensions. They slightly vary in the formulation of equations that are linearized and in the differentiability concepts that were employed. In the following presentation of our algorithm we take the active set perspective, the interpretation as generalized Newton method is discussed later. In the sequel we drop the index “ γ ” for the iterates.

Algorithm: (SSN)

- (1) Initialize $(\lambda^0, \xi^0, \mu^0, \mathbf{y}^0) \in L^2(\Gamma_c) \times L^2(\Gamma_c) \times L^2(\Gamma_c) \times \mathbf{Y}$, $\sigma > 0$ and set $k := 0$.
- (2) Determine the active and inactive sets

$$\begin{aligned} \mathcal{A}_c^{k+1} &= \{x \in \Gamma_c : \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}^k - d) > 0\}, \\ \mathcal{I}_c^{k+1} &= \Gamma_c \setminus \mathcal{A}_c^{k+1}, \\ \mathcal{A}_{f,-}^{k+1} &= \{x \in \Gamma_c : \xi^k + \sigma(\mu^k + \mathfrak{F}g) < 0\}, \\ \mathcal{A}_{f,+}^{k+1} &= \{x \in \Gamma_c : \xi^k + \sigma(\mu^k - \mathfrak{F}g) > 0\}, \\ \mathcal{I}_f^{k+1} &= \Gamma_c \setminus (\mathcal{A}_{f,-}^{k+1} \cup \mathcal{A}_{f,+}^{k+1}). \end{aligned}$$

- (3) If $k \geq 1$, $\mathcal{A}_c^{k+1} = \mathcal{A}_c^k$, $\mathcal{A}_{f,-}^{k+1} = \mathcal{A}_{f,-}^k$ and $\mathcal{A}_{f,+}^{k+1} = \mathcal{A}_{f,+}^k$ stop, else
- (4) Solve

$$\begin{aligned} a(\mathbf{y}^{k+1}, \mathbf{z}) - L(\mathbf{z}) + (\mu^{k+1}, \tau_T \mathbf{z})_{\Gamma_c} + (\lambda^{k+1}, \tau_N \mathbf{z})_{\Gamma_c} &= 0 \text{ for all } \mathbf{z} \in \mathbf{Y}, \\ \lambda^{k+1} &= 0 \text{ on } \mathcal{I}_c^{k+1}, \quad \lambda^{k+1} = \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}^{k+1} - d) \text{ on } \mathcal{A}_c^{k+1}, \\ \mu^{k+1} - \hat{\mu} - \gamma_2 \tau_T \mathbf{y}^{k+1} &= 0 \text{ on } \mathcal{I}_f^{k+1}, \\ \mu^{k+1} &= -\mathfrak{F}g \text{ on } \mathcal{A}_{f,-}^{k+1}, \quad \mu^{k+1} = \mathfrak{F}g \text{ on } \mathcal{A}_{f,+}^{k+1}. \end{aligned}$$

(5) Set

$$\xi^{k+1} := \begin{cases} \tau_T \mathbf{y}^{k+1} + \gamma_2^{-1}(\hat{\mu} + \mathfrak{F}g) & \text{on } \mathcal{A}_{f,-}^{k+1}, \\ \tau_T \mathbf{y}^{k+1} + \gamma_2^{-1}(\hat{\mu} - \mathfrak{F}g) & \text{on } \mathcal{A}_{f,+}^{k+1}, \\ 0 & \text{on } \mathcal{I}_f^{k+1}, \end{cases}$$

$k := k + 1$ and go to Step 2.

Note that there exists a unique solution to the system in Step 4, since it represents the necessary and sufficient optimality conditions for the equality constrained auxiliary problem

$$\min_{\substack{\lambda=0 \text{ on } \mathcal{I}_c^{k+1}, \\ \mu=-\mathfrak{F}g \text{ on } \mathcal{A}_{f,-}^{k+1}, \mu=\mathfrak{F}g \text{ on } \mathcal{A}_{f,+}^{k+1}}} J_{\gamma_1, \gamma_2}^*(\lambda, \mu),$$

with J_{γ_1, γ_2}^* as defined in (3.1) that clearly has a unique solution. One can prove the following lemma that justifies the stopping criterion in Step 3 of Algorithm (SSN).

Lemma 4.1. *If Algorithm (SSN) stops, the last iterate \mathbf{y}^k is the solution to the primal problem $(\mathcal{P}_{\gamma_1, \gamma_2})$ and (λ^k, μ^k) solves the dual problem $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$.*

Proof. The proof relies on the fact that, if the active sets coincide for two consecutive iterations, the iterates satisfy the complementarity conditions (3.3b) and (3.3c). \square

Provided we choose $\sigma = \gamma_2^{-1}$, the above algorithm can be interpreted as a semi-smooth Newton method in infinite-dimensional spaces. To show this assertion, we consider a reduced system instead of (3.3). Thereby, as in the dual problem $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$, the primal variable \mathbf{y} only acts as an auxiliary variable that is calculated from the dual variables (λ, μ) . We introduce the mapping $F : L^2(\Gamma_c) \times L^2(\Gamma_c) \rightarrow L^2(\Gamma_c) \times L^2(\Gamma_c)$ by

$$F(\lambda, \mu) = \begin{pmatrix} \lambda - \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_{\lambda, \mu} - d)) \\ \gamma_2 \tau_T \mathbf{y}_{\lambda, \mu} + \hat{\mu} - \mu - \max(0, \gamma_2 \tau_T \mathbf{y}_{\lambda, \mu} + \hat{\mu} - \mathfrak{F}g) \cdots \\ \cdots - \min(0, \gamma_2 \tau_T \mathbf{y}_{\lambda, \mu} + \hat{\mu} + \mathfrak{F}g) \end{pmatrix}, \tag{4.1}$$

where, for given λ and μ we denote by $\mathbf{y}_{\lambda, \mu}$ the solution to

$$a(\mathbf{y}, \mathbf{z}) - L(\mathbf{z}) + (\mu, \tau_T \mathbf{z})_{\Gamma_c} + (\lambda, \tau_N \mathbf{z})_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}. \tag{4.2}$$

Note that for $(\lambda, \mu) \in L^2(\Gamma_c) \times L^2(\Gamma_c)$ we have that $\tau_N \mathbf{y}_{\lambda, \mu}, \tau_T \mathbf{y}_{\lambda, \mu} \in H^{1/2}(\Gamma_c)$. Since $H^{1/2}(\Gamma_c)$ embeds continuously into $L^p(\Gamma_c)$ for every $p < \infty$, we have the following composition of mappings (for each $2 < p < \infty$):

$$\begin{cases} L^2(\Gamma_c) \times L^2(\Gamma_c) & \rightarrow & L^p(\Gamma_c) & \xrightarrow{\Theta} & L^2(\Gamma_c), \\ (\lambda, \mu) & \mapsto & \tau_N \mathbf{y}_{\lambda, \mu} & \mapsto & \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_{\lambda, \mu} - d)). \end{cases} \tag{4.3}$$

Since the mapping Θ involves a norm gap under the max-functional, it is Newton differentiable due to Theorem 1.2, which shows that the first component of F is Newton differentiable. A similar observation holds for the second component as well, and thus the whole mapping F is Newton differentiable. Hence, we can apply the semi-smooth Newton method to the equation $F(\lambda, \mu) = 0$. Calculating the explicit form of the Newton step leads to Algorithm (SSN) with $\sigma = \gamma_2^{-1}$. This close relationship between the primal-dual active set strategy and a specific application of semi-smooth Newton methods, first observed in [18], leads to the following local convergence result for Algorithm (SSN).

Theorem 4.2. *Suppose that there exists a constant $g_0 > 0$ with $\mathfrak{F}g \geq g_0$, further that $\sigma \geq \gamma_2^{-1}$ and that $\|\lambda^0 - \lambda_\gamma\|_{\Gamma_c}, \|\mu^0 - \mu_\gamma\|_{\Gamma_c}$ are sufficiently small. Then the iterates $(\lambda^k, \xi^k, \mu^k, \mathbf{y}^k)$ of (SSN) converge superlinearly to $(\lambda_\gamma, \xi_\gamma, \mu_\gamma, \mathbf{y}_\gamma)$ in $L^2(\Gamma_c) \times L^2(\Gamma_c) \times L^2(\Gamma_c) \times \mathbf{Y}$.*

Proof. The proof consists of two steps: first we prove the assertion for $\sigma = \gamma_2^{-1}$ and then we utilize this result for the general case $\sigma \geq \gamma_2^{-1}$.

Step 1. For $\sigma = \gamma_2^{-1}$ (SSN) is a semi-smooth Newton method for the equation $F(\lambda, \mu) = 0$ (F as defined in (4.1)). We already argued Newton differentiability of F . To apply Theorem 1.2, it remains to show that the generalized derivatives have uniformly bounded inverses, which is easy to verify (see, e.g., the proof for *Thm. 2.2* in [23] or of *Thm. 5.1* in [34]). Clearly, the superlinear convergence of (λ^k, μ^k) carries over to the variables ξ^k and \mathbf{y}^k .

Step 2. For $\sigma > \gamma_2^{-1}$ we cannot use the above argument directly. Nevertheless, one can prove superlinear convergence of the iterates by showing that in a neighborhood of the solution the iterates of (SSN) with $\sigma > \gamma_2^{-1}$ coincide with those of (SSN) with $\sigma = \gamma_2^{-1}$. The argumentation for this fact exploits the smoothing properties of the Neumann-to-Dirichlet mapping for the elasticity equation: first, we again consider the case $\sigma = \gamma_2^{-1}$. Clearly, for all $k \geq 1$ we have $\lambda^k - \lambda^{k-1} \in L^2(\Gamma_c)$ and $\mu^k - \mu^{k-1} \in L^2(\Gamma_c)$. The corresponding difference $\mathbf{y}^k - \mathbf{y}^{k-1}$ of the primal variables satisfies

$$a(\mathbf{y}^k - \mathbf{y}^{k-1}, \mathbf{z}) + (\mu^k - \mu^{k-1}, \tau_T \mathbf{z})_{\Gamma_c} + (\lambda^k - \lambda^{k-1}, \tau_N \mathbf{z})_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}.$$

From regularity results for elliptic variational equalities (see [3]) it follows that there exists a constant $C > 0$ such that

$$\|\tau_T \mathbf{y}^k - \tau_T \mathbf{y}^{k-1}\|_{C^0(\Gamma_c)} \leq C(\|\lambda^k - \lambda^{k-1}\|_{\Gamma_c} + \|\mu^k - \mu^{k-1}\|_{\Gamma_c}). \tag{4.4}$$

We now show that (4.4) implies

$$\mathcal{A}_{f,-}^k \cap \mathcal{A}_{f,+}^{k+1} = \mathcal{A}_{f,+}^k \cap \mathcal{A}_{f,-}^{k+1} = \emptyset \tag{4.5}$$

provided that $\|\lambda^0 - \lambda_\gamma\|_{\Gamma_c}$ and $\|\mu^0 - \mu_\gamma\|_{\Gamma_c}$ are sufficiently small. If $\mathcal{B} := \mathcal{A}_{f,-}^k \cap \mathcal{A}_{f,+}^{k+1} \neq \emptyset$, then it follows that $\tau_T \mathbf{y}^{k-1} + \gamma_2^{-1}(\hat{\mu} + \mathfrak{F}g) < 0$ and $\tau_T \mathbf{y}^k + \gamma_2^{-1}(\hat{\mu} - \mathfrak{F}g) > 0$ on \mathcal{B} , which implies that $\tau_T \mathbf{y}^k - \tau_T \mathbf{y}^{k-1} > 2\gamma_2^{-1}\mathfrak{F}g \geq 2\gamma_2^{-1}\mathfrak{F}g_0 > 0$ on \mathcal{B} . This contradicts (4.4) provided that $\|\lambda^0 - \lambda_\gamma\|_{\Gamma_c}$ and $\|\mu^0 - \mu_\gamma\|_{\Gamma_c}$ are sufficiently small. Analogously, one can show that $\mathcal{A}_{f,+}^k \cap \mathcal{A}_{f,-}^{k+1} = \emptyset$.

We now choose an arbitrary $\sigma \geq \gamma_2^{-1}$ and assume that (4.5) holds for (SSN) if σ was chosen γ_2^{-1} . Then we can argue that in a neighborhood of the solution the iterates of (SSN) are independent of $\sigma \geq \gamma_2^{-1}$. To verify this assertion we separately consider the sets $\mathcal{I}_f^k, \mathcal{A}_{f,-}^k$ and $\mathcal{A}_{f,+}^k$. On \mathcal{I}_f^k we have that $\xi^k = 0$ and thus σ has no influence when determining the new active and inactive sets. On the set $\mathcal{A}_{f,-}^k$ we have $\mu^k = -\mathfrak{F}g$. Here, we consider two types of sets: Firstly, sets where $\xi^k < 0$ belong to $\mathcal{A}_{f,-}^{k+1}$ for the next iteration independently from σ . And, secondly, if $\xi^k > 0$ we have

$$\xi^k + \sigma(\mu^k - \mathfrak{F}g) = \xi^k - 2\sigma\mathfrak{F}g.$$

Sets where $\xi^k - 2\sigma\mathfrak{F}g \leq 0$ are transferred to \mathcal{I}_f^{k+1} , and those where $0 < \xi^k - 2\sigma\mathfrak{F}g \leq \xi^k - 2\gamma_2^{-1}\mathfrak{F}g$ belong to $\mathcal{A}_{f,+}^{k+1}$ for the next iteration. However, the case that $x \in \mathcal{A}_{f,-}^k \cap \mathcal{A}_{f,+}^{k+1}$ cannot occur for $\sigma \geq \gamma_2^{-1}$, since it is already ruled out by (4.5) for $\sigma = \gamma_2^{-1}$.

This shows that in a neighborhood of the solution the iterates are the same for all $\sigma \geq \gamma_2^{-1}$, and thus the superlinear convergence result from Step 1 carries over to the general case $\sigma \geq \gamma_2^{-1}$, which ends the proof. \square

Aside from the assumption that $\|\lambda^0 - \lambda_\gamma\|_{\Gamma_c}$ and $\|\mu^0 - \mu_\gamma\|_{\Gamma_c}$ are sufficiently small, σ controls the probability that points are moved from the lower active set to the upper or vice versa in one iteration. Smaller values for σ make it more likely that points belong to $\mathcal{A}_{f,-}^k \cap \mathcal{A}_{f,+}^{k+1}$ or $\mathcal{A}_{f,+}^k \cap \mathcal{A}_{f,-}^{k+1}$. In the numerical realization of (SSN) it

turns out that choosing small values for σ may not be optimal, since this possibly leads to the following behavior: Points that are active with respect to the upper bound become active with respect to the lower bound in the next iteration, and vice versa. This in turn may lead to cycling of the iterates. Such undesired behavior can be overcome by choosing larger values for σ , e.g., $\sigma = 1$.

5. AUGMENTED LAGRANGIAN METHOD

There are several possibilities to utilize a sequence of solutions of $(\mathcal{P}_{\gamma_1, \gamma_2})$ to obtain a solution of (\mathcal{P}) . Clearly, one approach consists in letting the regularization parameters tend to infinity. Then, Theorem 3.1 guarantees that the regularized solution variables converge to the solution of the original problem. A different approach utilizes the first-order augmented Lagrangian method that is based on an update strategy for the dual variables. This method was introduced for discretized contact and friction problems in [1]. It is related to the Uzawa algorithm, a commonly used algorithm for the treatment of saddle point problems. More precisely, the augmented Lagrangian method can be interpreted as implicit version of the Uzawa method (cp., [22]). The main advantage of the augmented Lagrangian method compared to the the Uzawa method is its unconditional convergence for all penalty (or regularization) parameters $\gamma_1, \gamma_2 > 0$, whereas Uzawa’s method only converges conditionally, requiring smallness conditions for the parameters γ_1, γ_2 , which may lead to very slow convergence. The penalty as well as the first order augmented Lagrangian methods require to solve a nonlinear problem in every iteration step compared to the linear auxiliary problem which need to be solved in the Uzawa algorithm. These nonlinear problems are exactly of the form $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$ and we can therefore use the strategies presented in the previous sections for the solution of the auxiliary problems. The augmented Lagrangian approach applies for the solution of contact problems with given friction in infinite dimensions provided that the solution variables $(\bar{\lambda}, \bar{\mu}) \in L^2(\Gamma_c) \times L^2(\Gamma_c)$. As shown in Section 2.2, $g \in L^2(\Gamma_c)$ implies L^2 -regularity for $\bar{\mu}$. Assuming that this regularity holds, we present an augmented Lagrangian method for the solution of (\mathcal{P}) and (\mathcal{P}^*) . We now specify the algorithm.

Algorithm: (ALM)

- (1) Choose $(\lambda^0, \mu^0) \in L^2(\Gamma_c) \times L^2(\Gamma_c)$ and set $l := 0$.
- (2) Choose $\hat{\gamma}_1^{l+1}, \hat{\gamma}_2^{l+1} > 0$, and solve $(\mathcal{P}_{\hat{\gamma}_1, \hat{\gamma}_2})$ with $\hat{\lambda} := \lambda^l$ and $\hat{\mu} = \mu^l$, i.e., determine $(\mathbf{y}^{l+1}, \lambda^{l+1}, \mu^{l+1}) \in \mathbf{Y} \times L^2(\Gamma_c) \times L^2(\Gamma_c)$ such that

$$\begin{aligned}
 &a(\mathbf{y}^{l+1}, \mathbf{z}) - L(\mathbf{z}) + (\lambda^{l+1}, \tau_N \mathbf{z})_{\Gamma_c} + (\mu^{l+1}, \tau_T \mathbf{z})_{\Gamma_c} = 0 \text{ for all } \mathbf{z} \in \mathbf{Y}, \\
 &\lambda^{l+1} - \max(0, \lambda^l + \hat{\gamma}_1^{l+1}(\tau_N \mathbf{y}^{l+1} - d)) = 0 \text{ on } \Gamma_c, \\
 &\gamma_2 \tau_T \mathbf{y}^{l+1} + \mu^l - \mu^{l+1} - \max(0, \gamma_2 \tau_T \mathbf{y}^{l+1} + \mu^l - \mathfrak{F}g) - \min(0, \gamma_2 \tau_T \mathbf{y}^{l+1} + \mu^l + \mathfrak{F}g) = 0 \text{ on } \Gamma_c.
 \end{aligned}$$

- (3) Update $l := l + 1$ and go to Step 2.

In the theorem below we denote by $\bar{\mathbf{y}}$ and $(\bar{\lambda}, \bar{\mu})$ the solutions of (\mathcal{P}) and (\mathcal{P}^*) , respectively. Then the following global convergence result (i.e. convergence from arbitrary initialization) holds true.

Theorem 5.1. *For every choice of parameters $0 < \gamma_1^0 \leq \gamma_1^1 \leq \gamma_1^2 \leq \dots$ and $0 < \gamma_2^0 \leq \gamma_2^1 \leq \gamma_2^2 \leq \dots$ the iterates (λ^l, μ^l) of (ALM) converge weakly to $(\bar{\lambda}, \bar{\mu})$ in $L^2(\Gamma_c) \times L^2(\Gamma_c)$. Furthermore, the corresponding iterates \mathbf{y}^l converge strongly to $\bar{\mathbf{y}}$ in \mathbf{Y} .*

Proof. The proof is given in Appendix B. □

Let us briefly comment on the role of the parameters γ_1^l, γ_2^l in (ALM). One may start with moderate values in Step 2 of (ALM) and increase these values during the iteration. However, the iterates of (ALM) converge without requiring that γ_1^l, γ_2^l tend to infinity which is in marked contrast to pure penalty methods.

6. THE CONTACT PROBLEM WITH COULOMB FRICTION

In this section we discuss a regularized Coulomb friction problem and prove existence of a solution for this problem. Possible generalizations of the approach for the Coulomb friction problem without regularization are discussed. Then we propose two fixed point algorithms for the numerical realization of the Coulomb friction problem that make use of solutions of the contact problem with Tresca friction.

6.1. Regularized Coulomb friction problem

The regularization for the Signorini contact problem with Coulomb friction corresponds to the regularization in $(\mathcal{P}_{\gamma_1\gamma_2})$ and $(\mathcal{P}_{\gamma_1\gamma_2}^*)$ for the problem with given friction. The variational formulation of the problem to be considered is

$$\begin{aligned}
 & a(\mathbf{y}, \mathbf{z} - \mathbf{y}) + (\max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y} - d)), \tau_N(\mathbf{z} - \mathbf{y}))_{\Gamma_c} - L(\mathbf{z} - \mathbf{y}) \\
 & + \frac{1}{\gamma_2} \int_{\Gamma_c} \mathfrak{F} \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y} - d)) \{h(\tau_T \mathbf{z}, \hat{\mu}) - h(\tau_T \mathbf{y}, \hat{\mu})\} \, dx \geq 0
 \end{aligned} \tag{6.1}$$

for all $\mathbf{z} \in \mathbf{Y}$, with $h(\cdot, \cdot)$ as defined in (3.2). Similarly as in Section 2.3 for the original Coulomb friction problem, one can obtain the solution to (6.1) by means of a sequence of regularized Tresca friction problems. For this purpose we introduce the cone of nonnegative L^2 -functions as $L^2_+(\Gamma_c) := \{\xi \in L^2(\Gamma_c) : \xi \geq 0 \text{ a.e.}\}$ and the mapping $\Psi_\gamma : L^2_+(\Gamma_c) \rightarrow L^2_+(\Gamma_c)$ defined by $\Psi_\gamma(g) := \lambda_\gamma$, where λ_γ is given by $\lambda_\gamma = \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_\gamma - d))$ with \mathbf{y}_γ denoting the unique solution of the regularized contact problem with friction $g \in L^2_+(\Gamma_c)$. We now show that the regularized problem with Coulomb friction (6.1) admits a solution using the fixed point characterization. In a first step, we investigate the mapping $\Phi_\gamma : L^2_+(\Gamma_c) \rightarrow \mathbf{Y}$ that maps a given friction $g \in L^2_+(\Gamma_c)$ to the corresponding solution \mathbf{y}_γ of $(\mathcal{P}_{\gamma_1\gamma_2})$. In the next lemma we state that Φ_γ is Lipschitz-continuous. For the technical proof we refer to [33].

Lemma 6.1. *For every $\gamma_1, \gamma_2 > 0$ and $\hat{\lambda} \in L^2(\Gamma_c), \hat{\mu} \in L^2(\Gamma_c)$ the mapping Φ_γ defined above is Lipschitz-continuous with constant*

$$\mathfrak{L} = \frac{\|\mathfrak{F}\|_\infty c_1}{\kappa}, \tag{6.2}$$

where $\|\mathfrak{F}\|_\infty$ denotes the essential supremum of \mathfrak{F} , κ the coercivity constant of $a(\cdot, \cdot)$ and c_1 the continuity constant of the trace mapping from \mathbf{Y} to $L^2(\Gamma_c)$. In particular, the Lipschitz constant \mathfrak{L} does not depend on the regularization parameters γ_1, γ_2 .

We now address properties of the mapping Ψ_γ .

Lemma 6.2. *For every $\gamma_1, \gamma_2 > 0$ and $\hat{\lambda} \in L^2(\Gamma_c), \hat{\mu} \in L^2(\Gamma_c)$ the mapping $\Psi_\gamma : L^2_+(\Gamma_c) \rightarrow L^2_+(\Gamma_c)$ is compact and Lipschitz-continuous with constant*

$$\mathfrak{L} = \frac{c\gamma}{\kappa} \|\mathfrak{F}\|_\infty,$$

where c is a constant resulting from trace theorems.

Proof. We consider the following composition of mappings:

$$\begin{array}{ccccccc}
 L^2_+(\Gamma_c) & \xrightarrow{\Phi_\gamma} & \mathbf{Y} & \xrightarrow{\Theta} & L^2(\Gamma_c) & \xrightarrow{\Upsilon} & L^2_+(\Gamma_c), \\
 g & \mapsto & \mathbf{y} & \mapsto & \tau_N \mathbf{y} & \mapsto & \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y} - d)).
 \end{array} \tag{6.3}$$

From Lemma 6.1 we already know that Φ_γ is Lipschitz-continuous. The mapping Θ consists of the linear trace mapping from \mathbf{Y} into $H^{1/2}(\Gamma_c)$ and the compact embedding of this space into $L^2(\Gamma_c)$. Therefore, it is compact and linear, in particular Lipschitz-continuous with a constant we denote by $c_2 > 0$. Finally, since

$$\|\max(0, \hat{\lambda} + \gamma_1(\xi - d)) - \max(0, \hat{\lambda} + \gamma_1(\tilde{\xi} - d))\|_{\Gamma_c} \leq \gamma_1 \|\xi - \tilde{\xi}\|_{\Gamma_c} \tag{6.4}$$

for all $\xi, \tilde{\xi} \in L^2(\Gamma_c)$, the mapping Υ is Lipschitz-continuous with constant γ_1 . From the fact that Ψ_γ is the composition of the mappings $\Upsilon, \Theta, \Phi_\gamma$, namely $\Psi_\gamma = \Upsilon \circ \Theta \circ \Phi_\gamma$, we conclude that Ψ_γ is Lipschitz-continuous with constant

$$\mathfrak{L} = \frac{c_1 c_2 \gamma_1}{\kappa} \|\mathfrak{F}\|_\infty, \tag{6.5}$$

where c_1, c_2 are constants from trace theorems. Concerning the compactness, we clearly have that the composition of Θ and Φ_γ is compact. It remains to show that L^2 -convergent sequences remain L^2 -convergent under the mapping Υ . This follows immediately from (6.4), which ends the proof. \square

We can now show that the regularized contact problem with Coulomb friction has a solution.

Theorem 6.3. *The mapping Ψ_γ admits at least one fixed point, i.e., the regularized Coulomb friction problem (6.1) admits a solution. If $\|\mathfrak{F}\|_\infty$ is such that \mathfrak{L} as defined in (6.5) is smaller than 1, the solution is unique.*

Proof. We apply the Leray-Schauder fixed point theorem (see [11], p. 222) to the mapping $\Psi_\gamma : L^2(\Gamma_c) \rightarrow L^2(\Gamma_c)$. Using Lemma 6.2, it suffices to show that λ is bounded in $L^2(\Gamma_c)$ independently of g . This is clear taking into account the dual problem $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$. Indeed,

$$\min_{\lambda \geq 0, |\mu| \leq \mathfrak{F}g \text{ a.e. on } \Gamma_c} J_{\gamma_1, \gamma_2}^*(\lambda, \mu) \leq \min_{\lambda \geq 0} J_{\gamma_1, \gamma_2}^*(\lambda, 0) < \infty.$$

Hence, the Leray-Schauder theorem guarantees the existence of a solution to the regularized Coulomb friction problem. Uniqueness of the solution holds if \mathfrak{F} is such that \mathfrak{L} as defined in (6.5) is smaller than 1, due to the fact that then Ψ_γ is a contraction. \square

6.2. Remarks on the Coulomb friction problem without regularization

Here we briefly discuss the question, which methods utilized in the previous section also apply in the case of Coulomb friction without regularization and which do not. Firstly, we are interested in an analogue to Lemma 6.1 in the case without regularization. As observed above, the Lipschitz-constant in Lemma 6.1 is independent of γ_1, γ_2 , which suggests that the result does not require any regularizing term. To verify this conjecture we define the mapping $\Phi : L_+^2(\Gamma_c) \rightarrow \mathbf{Y}$ that maps a given friction $g \in L_+^2(\Gamma_c)$ to the corresponding solution \mathbf{y} of (\mathcal{P}) . Lipschitz-continuity of Ψ is stated in the next lemma, for its short proof we refer to [33].

Lemma 6.4. *The mapping Φ defined above is Lipschitz-continuous with constant*

$$\mathfrak{L} = \frac{\|\mathfrak{F}\|_\infty c_1}{\kappa},$$

where $\|\mathfrak{F}\|_\infty$ denotes the essential supremum of \mathfrak{F} , κ the coercivity constant of $a(\cdot, \cdot)$ and c_1 a constant from a trace theorem.

We can now discuss, which parts of Lemma 6.4 and Theorem 6.3 require a regularization and which do not. Obviously, the mapping Θ appearing in the proof of Lemma 6.2 does not involve any regularization. The mapping Υ is Lipschitz-continuous with constant γ_1 , which indicates that the regularization of the contact condition is required to obtain Lipschitz-continuity. Thus, for Lemma 6.2 the regularization with respect to the contact condition is essential. However, for proving existence of a solution to the Coulomb friction problem the regularization of the friction condition is not essential, since neither the mapping Φ nor Θ and Υ involve γ_2 . Using this observation and Lemma 6.4, one can verify that both Lemma 6.2 and Theorem 6.3 also hold true for the case that we apply a regularization only for the contact condition, while leaving the non-differentiability in (\mathcal{P}) unchanged.

6.3. Algorithm for the solution of the Coulomb friction problem

As mentioned before, the fixed point idea presented in Section 6.1 can be exploited numerically for the solution of the regularized Coulomb friction problem. This idea and slight modifications of this idea are commonly used to calculate the solution of the Coulomb friction problem by a sequence of Tresca friction problems (see, e.g., [16, 26, 27, 29]). The pure fixed point algorithm is given next.

Algorithm: (FP)

- (1) Choose $\gamma_1, \gamma_2 > 0$, $\hat{\lambda}$ and $\hat{\mu}$. Initialize $g^0 \in L^2(\Gamma_c)$, and set $m := 0$.
- (2) Determine the solution (λ^m, μ^m) to problem $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$ with given friction g^m .
- (3) Update $g^{m+1} := \lambda^m$, $m := m + 1$ and, unless an appropriate stopping criterion is met, go to Step 2.

Provided that $\|\mathfrak{F}\|_\infty$ is sufficiently small, Lemma 6.2 leads to a convergence result for the above algorithm.

Theorem 6.5. *Suppose that $\|\mathfrak{F}\|_\infty$ is sufficiently small, then Algorithm (FP) converges regardless of the initialization.*

Proof. The proof follows immediately from the fact that, provided $\|\mathfrak{F}\|_\infty$ is sufficiently small, the Lipschitz constant of the mapping Ψ_γ is smaller than 1. \square

One can consider various modifications of Algorithm (FP). In [16] a splitting type algorithm, based on a finite dimensional dual formulation of the Tresca friction problem is presented and its numerical performance is tested. In [26] a Gauß-Seidel-like generalization of (FP) in the framework of monotone multigrid methods is proposed, and the author reports on favorable numerical results. Here we propose a modification of (FP) that combines both the first order augmented Lagrangian update and the fixed point idea of Algorithm (FP).

Algorithm: (ALM-FP)

- (1) Initialize $\gamma_1^1, \gamma_2^1 > 0$, $(\hat{\lambda}^1, \hat{\mu}^1) \in L^2(\Gamma_c) \times L^2(\Gamma_c)$ and $g^0 \in L^2(\Gamma_c)$, $m := 0$.
- (2) Choose $\gamma_1^m, \gamma_2^m > 0$ and determine the solution (λ^m, μ^m) to problem $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$ with given friction g^m and $\hat{\lambda} := \hat{\lambda}^m, \hat{\mu} := \hat{\mu}^m$.
- (3) Update $g^{m+1} := \lambda^m$, $\hat{\lambda}^{m+1} := \lambda^m$, $\hat{\mu}^{m+1} := \mu^m$ and $m := m + 1$. Unless an appropriate stopping criterion is met, go to Step 2.

For the following brief discussion of the above algorithm, we assume that a solution to the Coulomb friction problem as defined in Section 2.3 exists and that the solution variables are sufficiently smooth. To be precise, we assume that the fixed point λ^* of the mapping Ψ (for its definition see p. 834) is in $L^2(\Gamma_c)$. Then, with the variables $\mathbf{y}^* \in \mathbf{Y}$ and $\mu^* \in L^2(\Gamma_c)$ corresponding to λ^* , we have for $\gamma_1, \gamma_2 > 0$ that

$$\begin{aligned} a(\mathbf{y}^*, \mathbf{z}) - L(\mathbf{z}) + (\lambda^*, \tau_N \mathbf{z})_{\Gamma_c} + (\mu^*, \tau_T \mathbf{z})_{\Gamma_c} &= 0 \text{ for all } \mathbf{z} \in \mathbf{Y}, \\ \lambda^* - \max(0, \lambda^* + \gamma_1(\tau_N \mathbf{y}^* - d)) &= 0 \text{ on } \Gamma_c, \\ \gamma_2 \tau_T \mathbf{y}^* - \max(0, \gamma_2 \tau_T \mathbf{y}^* + \mu^* - \mathfrak{F} \lambda^*) - \min(0, \gamma_2 \tau_T \mathbf{y}^* + \mu^* + \mathfrak{F} \lambda^*) &= 0 \text{ on } \Gamma_c. \end{aligned}$$

As can be seen easily, these conditions are also satisfied by a fixed point of (ALM-FP), i.e., provided convergence of (ALM-FP) the limit variables are a solution to the original, i.e., non-regularized contact problem with Coulomb friction.

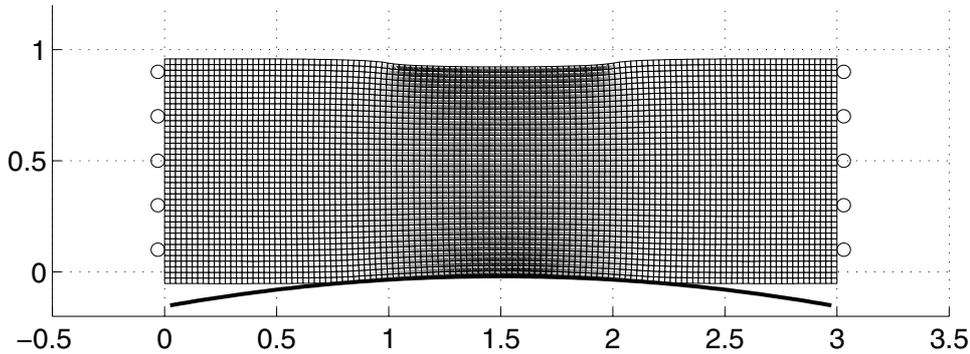


FIGURE 2. Example 1: deformed mesh (deformation multiplied by 20), gray tones visualize the elastic shear energy density, $\nu = 0.4$, $g \equiv 1.5$, $\gamma_1 = \gamma_2 = 10^8$.

7. NUMERICAL RESULTS

Our numerical implementation of the algorithms presented in the previous sections uses the MATLAB-code published in [2], that uses linear and bilinear finite elements for the discretization of the elasticity equations without contact and friction. Using (SSN), (ALM), (FP) and (ALM-FP) it has been modified such that it applies to problems involving contact and friction as well.

In the semi-smooth Newton method (SSN) always $\sigma = 1$ is chosen and, unless otherwise specified, the method is initialized with the solution of the unconstrained dual problem, *i.e.*, the solution of $(\mathcal{P}_{\gamma_1, \gamma_2}^*)$ neglecting the constraints. Unless otherwise stated $\hat{\lambda} = 0$, $\hat{\mu} = 0$ for (SSN) and the augmented Lagrangian method (ALM) is initialized with $\lambda^0 = 0$ and $\mu^0 = 0$. For the problems with given friction we always choose $\mathfrak{F} = 1$. The algorithms for the Coulomb friction problem (FP) and (ALM-FP) are initialized with the solution of the pure contact problem (and with $g^0 \equiv 0$). As in [26, 29] the outer (*i.e.*, the fixed point) iteration is terminated if

$$d_g^m := \frac{\|g^m - g^{m-1}\|_{\Gamma_c}}{\|g_m\|_{\Gamma_c}} \leq 10^{-7}.$$

7.1. Example 1

For this example the data are as follows: $\Omega = [0, 3] \times [0, 1]$, $\Gamma_c = [0, 3] \times \{0\}$, $\Gamma_n = [0, 3] \times \{1\}$ and the elastic body is subject to homogeneous Dirichlet conditions with respect to the horizontal displacement and homogeneous stress-free conditions with respect to the vertical displacement along $\{0\} \times [0, 1] \cup \{3\} \times [0, 1]$. Further, $\mathbf{f} = 0$ and

$$\mathbf{h} = \begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{on } [0, 1] \times \{1\} \cup [2, 3] \times \{1\}, \\ \begin{pmatrix} 0 \\ -20 \end{pmatrix} & \text{on } [1, 2] \times \{1\}. \end{cases}$$

For our test runs we choose $E = 5000$ and for the distance towards the rigid foundation $d(x_1) = 0.003(x_1 - 1.5)^2 + 0.001$. First we report on tests for $\nu = 0.4$ and given friction $g = 1.5$, then on tests with $\nu = 0.49$ and $g = 2.5$. The geometry of this example allows rigid motions of the complete body in vertical direction, since the elastic body is nowhere fixed with respect to the horizontal direction. Nevertheless, from the geometry it is clear that these motions are excluded if the elastic body is in contact with the foundation on some part of Γ_c .

We first summarize our testing of (SSN) for the problem with given friction $g \equiv 1.5$ and $\nu = 0.4$. The semi-smooth Newton algorithm detects the solution for $\gamma_1 = \gamma_2 = 10^8$ on a mesh of 120×40 finite elements after 9 iterations. The deformed mesh is depicted in Figure 2, where, for graphical presentation the deformation and the distance to the rigid foundation is multiplied by a factor of 20 and gray tones are utilized to visualize the

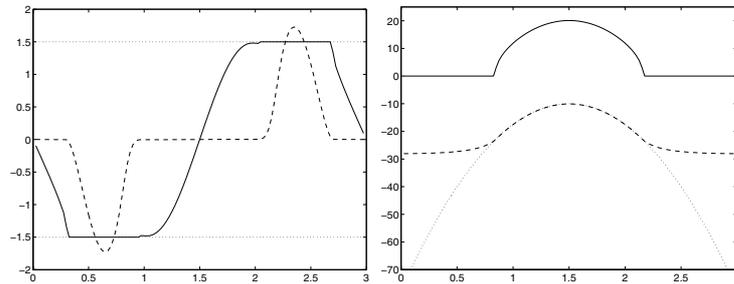


FIGURE 3. Example 1, left: dual variable μ_γ (solid) with bounds $\pm g$ (dotted) and tangential displacement $\tau_T \mathbf{y}_\gamma$ (multiplied by 10^4 , dashed). Right: multiplier λ_γ (solid), rigid foundation (multiplied by 10^4 , dotted) and normal displacement $\tau_N \mathbf{y}_\gamma$ (multiplied by 10^4 , dashed), $\nu = 0.4$, $g \equiv 1.5$, $\gamma_1 = \gamma_2 = 10^8$.

elastic shear energy density. The dual variable μ_γ , the corresponding bounds $\pm g$ and the tangential displacement are shown on the left of Figure 3. By inspection of the graphs one can verify that the complementarity conditions hold. Recall that active sets correspond to parts of the boundary, where the elastic body is sliding in tangential direction, while inactive sets correspond to sticky regions, *i.e.*, to sets where $\tau_T \mathbf{y}_\gamma = 0$. The main disadvantage of utilizing the Tresca friction law is that stick may occur on sets where the elastic body is *not* in contact with the foundation. This phenomenon also occurs for our example as can be seen from Figure 3, where tangential stick without contact occurs, *e.g.*, in the interval $[0, 0.3]$. This is clearly undesirable in modelling realistic physical phenomena, but it can occur in the Tresca friction model. This phenomenon does not cause any difficulties in the proposed solution strategies.

Let us now comment on the performance of the algorithm. To obtain convergence of the method independently of the initialization, it seems crucial to choose the parameter σ large enough, for instance, $\sigma = 1$. Setting $\sigma = \gamma_2^{-1}$ as suggested by the interpretation of the algorithm as infinite-dimensional semi-smooth Newton method may lead to cycling of the iterates unless the initialization is already close to the solution (see also the discussion at the end of Sect. 4). For $\sigma = 1$, the method converges for all initializations and for all choices of γ_1, γ_2 . We remark that the Newton-type method proposed in [1], which is similar to ours, does not use any additional parameter σ . Using our notation, this corresponds to the choice $\sigma = \gamma_2^{-1}$. This choice has the advantage that points are moved from the lower active set to the upper and vice versa within one iteration. On the other hand it can be the reason why the authors of [1] observe a cycling of the iterates (*cf.* [1], p. 369). In general it can be said that decreasing σ increases the likelihood that coordinates are moved from being active above to active below in within one iteration.

Since the elastic body is only fixed in horizontal direction, we cannot initialize the algorithm with the solution of the state equation (3.3a) for $\mu_\gamma = 0, \lambda_\gamma = 0$. Doing so would lead to a problem where both, friction and contact conditions are neglected and thus the body is nowhere fixed in horizontal direction. Therefore, as initialization we choose the solution of the problem without friction, but with forcing that $\tau_N \mathbf{y}_\gamma - d = 0$ in the interval $[1.25, 1.75]$. We tested several other initializations including $\tau_N \mathbf{y}_\gamma - d = 0$ on all of Γ_c , and found the algorithm to converge always after at most 12 iterations.

In the iteration process we usually observe a monotonicity of the active sets for the contact condition, namely, beginning from the second iteration, the new estimate for the active set is contained in the previous one. We do not observe a similar behavior for the active sets corresponding to the friction condition. Generally, it can be said that in this example the contact condition strongly dominates the convergence behavior.

From Table 1, which displays the number of iterations for various γ_1, γ_2 it can be seen that the number of iterations depends only weakly on γ_1, γ_2 . Next we investigate the speed of convergence of the iterates of (SSN).

TABLE 1. Example 1: number of iterations for different values of γ_1 and γ_2 , $\nu = 0.4$, $g \equiv 1.5$.

		γ_2					
		10^4	10^5	10^6	10^7	10^8	10^9
γ_1	10^4	6	7	8	8	8	8
	10^5	6	8	8	9	9	9
	10^6	7	8	12	12	13	13
	10^7	7	8	10	11	11	11
	10^8	7	8	10	11	11	11
	10^9	7	8	10	11	11	11

TABLE 2. Example 1: values for q^k , $\nu = 0.4$, $g \equiv 1.5$, $\gamma_1 = \gamma_2 = 10^4$.

k	1	2	3	4
q^k	1.11e-0	8.64e-1	3.00e-1	0

TABLE 3. Example 1: number of iterations on different grids and for various regularization parameters for symmetric initialization and non-symmetric initialization (number in parentheses), $\nu = 0.49$, $g \equiv 2.5$.

		$\gamma_1 = \gamma_2$				
grid		10^5	10^6	10^7	10^8	10^9
60×20		6(8)	7(9)	9(12)	9(12)	9(12)
120×40		7(8)	9(11)	14(16)	13(18)	13(18)
240×80		7(9)	9(11)	16(18)	17(21)	17(21)

For this purpose we report, for $k = 1, 2, \dots$ on the discrete analogues of

$$q^k := \frac{a(\mathbf{y}^{k+1} - \mathbf{y}_\gamma, \mathbf{y}^{k+1} - \mathbf{y}_\gamma)^{\frac{1}{2}}}{a(\mathbf{y}^k - \mathbf{y}_\gamma, \mathbf{y}^k - \mathbf{y}_\gamma)^{\frac{1}{2}}} + \frac{\|\lambda^{k+1} - \lambda_\gamma\|_{\Gamma_c}}{\|\lambda^k - \lambda_\gamma\|_{\Gamma_c}} + \frac{\|\mu^{k+1} - \mu_\gamma\|_{\Gamma_c}}{\|\mu^k - \mu_\gamma\|_{\Gamma_c}},$$

where $(\mathbf{y}_\gamma, \lambda_\gamma, \mu_\gamma)$ denote the solution variables and $(\mathbf{y}^k, \lambda^k, \mu^k)$ the iterates. The results are shown in Table 2, where we observe that q^k is monotonously decreasing indicating superlinear convergence of the iterates.

We now report on the number of iterations of (SSN) for differently fine discretizations. We consider the case $g \equiv 2.5$ and $\nu = 0.49$ and show in Table 3 the number of iterations for various regularization parameters on three different grids. Thereby, the first number corresponds to the standard initialization of the algorithm, *i.e.*, setting $\tau_N \mathbf{y}^0 = d$ on $[1.25, 1.75]$. The number in the parentheses corresponds to the non-symmetric initialization $\tau_N \mathbf{y}^0 = d$ on $[0, 0.5]$. One observes that the algorithm behaves only moderately mesh-dependent and that for all grids the number of iterations increases as γ_1, γ_2 increase. For $\gamma_1 = \gamma_2 = 10^5, 10^6$ we can observe an almost mesh-independent behavior. A possible explanation for this remarkable result is that for small regularization parameters the convergence region of the continuous method is large, such that we can observe mesh-independence of the semi-smooth Newton method (as analyzed in [17]).

The above results motivate the application of a continuation procedure with respect to the regularization parameter to reduce the overall number of iterations. In this strategy we calculate the solution for moderately large regularization parameters and utilize the obtained solution as initialization for larger regularization parameters. The outcome for our problem is reported in Table 4. We observe that continuation with respect to γ_1 and γ_2 can be used to reduce the overall number of iterations especially on fine grids. Moreover, one observes that the number of iterations is even more reduced for the case that a non-symmetric initialization for (SSN) is used.

TABLE 4. Example 1: number of iterations on different grids using continuation w.r. to γ_1, γ_2 (2nd and 3th column), the resulting overall number of iterations (\sum) and the number without continuation strategy (last column). The numbers in parentheses correspond to a non-symmetric initialization, $\nu = 0.49, g \equiv 2.5$.

grid	$\gamma_1 = \gamma_2$		$\gamma_1 = \gamma_2$
	10^5	$\rightarrow 10^9$	\sum
60×20	6(8)	+4	10(11)
120×40	7(8)	+5	12(13)
240×80	7(9)	+7	14(16)

TABLE 5. Example 1: tests for (ALM) with inexact solve of the auxiliary problem, $\#iter$ denotes the number of inner iteration, $\nu = 0.49, g \equiv 2.5$.

	l	1	2	3	4	5	6
$\gamma_1 = \gamma_2 = 10^6$	$\#iter_{SSN}$	4	4	3	3	2	3
	$d_{\lambda,\mu}^m$	6.85e-1	1.45e-1	2.73e-2	1.28e-2	7.85e-3	5.24e-3
$\gamma_1 = \gamma_2 = 10^7$	$\#iter_{SSN}$	4	4	4	4	2	2
	$d_{\lambda,\mu}^m$	7.27e-1	1.80e-1	5.26e-2	1.52e-2	1.74e-3	3.39e-4
$\gamma_1 = \gamma_2 = 10^{4+l}$	$\#iter_{SSN}$	4	4	4	2	2	
	$d_{\lambda,\mu}^m$	5.22e-1	1.43e-1	2.23e-3	3.15e-5	8.64e-8	

Finally, we report on tests of the augmented Lagrangian method (ALM), where we terminate the (ALM)-iteration as soon as

$$d_{\lambda,\mu}^l := \|\bar{\lambda} - \lambda^l\|_{\Gamma_c} + \|\bar{\mu} - \mu^l\|_{\Gamma_c} < 10^{-4}, \tag{7.1}$$

where for $\bar{\lambda}$ and $\bar{\mu}$ we take the solution of the regularized problem with $\gamma_1 = \gamma_2 = 10^{10}$. Each iteration step of (ALM) is initialized with the solution variables of the previous iteration. Applying the method with $\gamma_1 = \gamma_2 = 10^7$, 2 (ALM)-iterations and overall 19 iterations of (SSN) are needed to obtain $d_{\lambda,\mu} = 6.23e-5$. To accelerate the convergence, one can increase the regularization parameters by a factor of, e.g., 10 in every (ALM)-iteration. Starting with $\gamma_1 = \gamma_2 = 10^6$, the algorithm terminates after 2 (ALM)-iterations and overall 16 linear solves with $d_{\lambda,\mu}^2 = 2.71e-5$. We now turn to an inexact version of the above method, namely we allow a maximum of 4 (SSN)-iterations in the inner loop of (ALM), before we update λ and μ . Applying this inexact strategy for the inner problem in (ALM) may lead to iterates λ^l, μ^l that do not satisfy $\lambda^l \geq 0$ and $|\mu^l| \leq g$. Thus, in our inexact strategy, we apply a projection of the dual variables in order to make them feasible, before we perform the augmented Lagrangian update. The results for the resulting test runs are summarized in Table 5, where we also document the values for $d_{\lambda,\mu}^l$. It can be observed that for the first iterates of (ALM) with inexact solve of the auxiliary problem smaller regularization parameters lead to better results, which is due to the smaller ill-conditioning of the auxiliary problem. Thus, additionally to the augmented Lagrangian update we increase γ_1 and γ_2 , which leads to a fast convergence to the solution of the non-regularized contact problem with Tresca friction, see the last row of Table 5.

Next we turn to the problem with Coulomb friction, where we consider the case $\nu = 0.4$. The solution variables for $\mathfrak{F} = 0.3$ obtained with (ALM-FP) are shown on the left and in the middle of Figure 4. Note that the bounds for the variable μ_γ in Coulomb's friction problem are given by $\mathfrak{F}\lambda_\gamma$, meaning that these bounds are not given *a priori* as in the case of Tresca friction. The convergence behavior of (ALM-FP) is shown in the table in Figure 4.

In our tests we always observe convergence of both (FP) and (ALM-FP). Recall that, provided the solution is sufficiently regular, the iterates of (ALM-FP) converge to the solution of the original Coulomb friction problem, while the iterates of (FP) converge to the solution of the regularized Coulomb friction problem. Both strategies

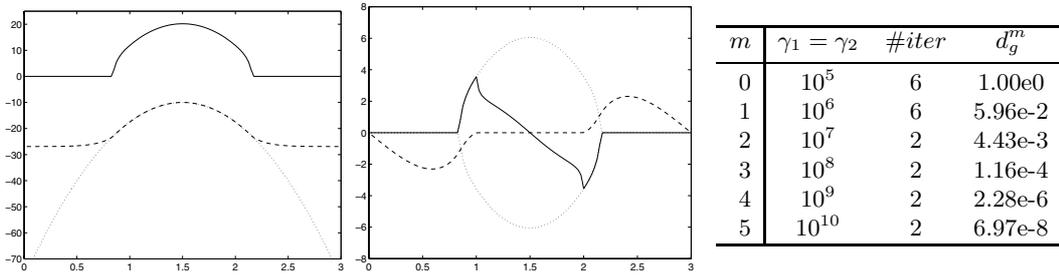


FIGURE 4. Example 1, left: multiplier λ_γ (solid), rigid foundation (multiplied by 10^4 , dotted) and normal displacement $\tau_N y_\gamma$ (multiplied by 10^4 , dashed), middle: variable μ_γ (solid) with bounds $\pm 0.3\lambda_\gamma$ (dotted) and tangential displacement y_γ (multiplied by 10^4 , dashed), right: number of inner iterations $\#iter$ for $\gamma_1 = \gamma_2$ for each outer iteration m of (ALM-FP), $\nu = 0.4$.

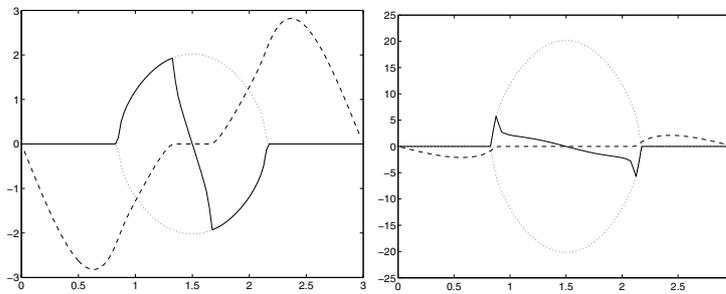


FIGURE 5. Example 1 with $\nu = 0.4$, left: variables μ_γ (solid) with bounds $\mathfrak{F}\lambda_\gamma$ (dotted) and tangential displacement y_γ (multiplied by 10^4 , dashed) $\mathfrak{F} = 0.1$. Right: same as left, but with $\mathfrak{F} = 1$.

require overall about 20 to 25 linear solves to detect the solution, that is, the computational effort for solving the Coulomb friction problem is comparable to the effort for the Tresca friction problem. Finally, in Figure 5 we show results obtained by (ALM-FP) for two different friction coefficients, namely for $\mathfrak{F} = 0.1$ and $\mathfrak{F} = 1$.

7.2. Example 2

In the previous example we could only observe a minor influence of frictional behavior onto the deformation, in particular onto the actual contact zone. This example has been constructed in order to investigate the convergence behavior of the algorithms for examples where the given friction and the friction coefficient significantly influences the deformation of the elastic body. We choose the same geometry as for Example 1, but the gap function $d = \max(0.0015, 0.003(x_1 - 1.5)^2 + 0.001)$ and $E = 10\,000$, $\nu = 0.45$. Further, we do not apply a traction force, rather we prescribe a nonzero deformation along the Dirichlet part of the boundary. As before, the zone of possible contact and friction is $\Gamma_c := [0, 3] \times \{0\}$, further on $\Gamma_n := [0, 3] \times \{1\} \cup \{0\} \times [0, 0.2] \cup \{3\} \times [0, 0.2]$ we assume traction free boundary conditions. On $\Gamma_d := \{0\} \times [0.2, 1] \cup \{3\} \times [0.2, 1]$ we prescribe the deformation as follows:

$$\tau \mathbf{y} = \begin{cases} \begin{pmatrix} 0.003(1-x_2) \\ -0.004 \end{pmatrix} & \text{on } \{0\} \times [0.2, 1], \\ \begin{pmatrix} -0.003(1-x_2) \\ -0.004 \end{pmatrix} & \text{on } \{3\} \times [0.2, 1]. \end{cases}$$

For $\gamma_1 = \gamma_2 = 10^8$ the semi-smooth Newton method detects the solution for given friction $g \equiv 1$ after 7 iterations. The corresponding deformed mesh and the elastic shear energy density are shown in Figure 6. As expected, in a neighborhood of the points $(0, 0.2)$ and $(3, 0.2)$, *i.e.*, the points where the boundary conditions change from

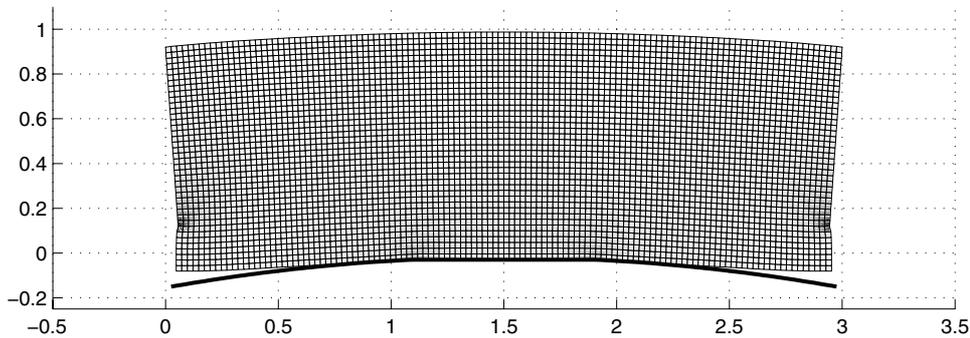


FIGURE 6. Example 2: Deformed mesh for $g \equiv 1$, gray tones visualize the elastic shear energy density.

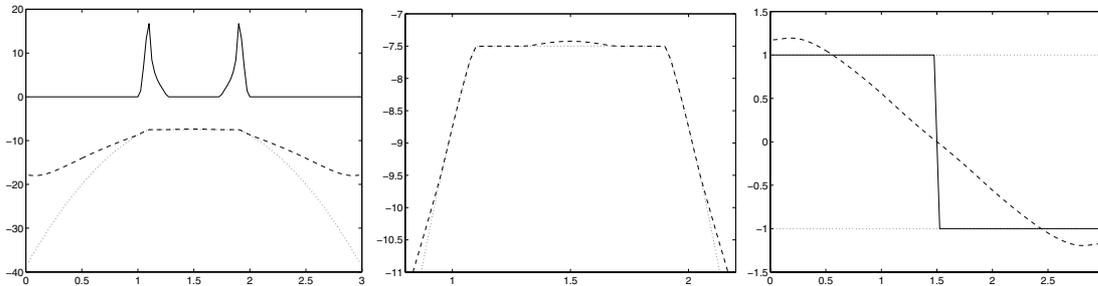


FIGURE 7. Example 2, left: multiplier λ_γ (solid), rigid foundation (multiplied by 5×10^3 , dotted) and normal displacement $\tau_N \mathbf{y}_\gamma$ (multiplied by 5×10^3 , dashed), Middle: magnification of detail of the left plot, Right figure: dual variable μ_γ (solid) with bounds $\pm \xi g$ (dotted) and tangential displacement \mathbf{y}_γ (multiplied by 5×10^3 , dashed).

Neumann to Dirichlet, we observe a stress concentration due to a local singularity of the solution. Further, we also observe a (small) stress concentration close to the points where the rigid foundation has the kinks.

On the left hand side of Figure 7 we depict the rigid foundation, the normal displacement and the corresponding multiplier. Magnifying the contact zone (see the plot in the middle of Fig. 7), one can observe that the body is not in contact with the rigid foundation in the interval $[1.3, 1.7]$, *i.e.*, in this interval the constraint on the normal deformation is inactive, which is also reflected in the fact that the corresponding dual variable λ_γ is zero. On the right hand side of Figure 7 we show the tangential displacement, the multiplier μ_γ and the bounds $\pm g$.

Next we investigate the influence of the given friction onto the deformation of the elastic body. Note that g directly influences μ_γ , while its influence on $\tau_N \mathbf{y}_\gamma$ and λ_γ is only due to the connection of the variables by means of the elasticity equation. The normal displacement and the corresponding multipliers λ_γ for $g \equiv 0, 1, 5$ are compared in Figure 8. For all values of g the algorithm finds the solution after 7 or 8 iterations, which shows a remarkable stability of (SSN) with respect to the bounds $\pm g$ for the dual variable μ_γ .

We now turn to the Coulomb friction problem and investigate the performance of (FP) and (ALM-FP). In Figure 9, we depict the normal and the tangential displacement with corresponding multipliers for $\mathfrak{F} = 2, 5, 10$. One observes that the friction coefficient significantly influences the deformation. For instance, in the case $\mathfrak{F} = 2$ the elastic body is in contact with the foundation in the interval $[1.4, 1.6]$, but it is not for $\mathfrak{F} = 5$ and $\mathfrak{F} = 10$. The solutions are obtained using (FP) and (ALM-FP). The methods perform comparably well and require overall between 20 and 25 linear solves to stop with $d_g^m \leq 10^{-7}$.

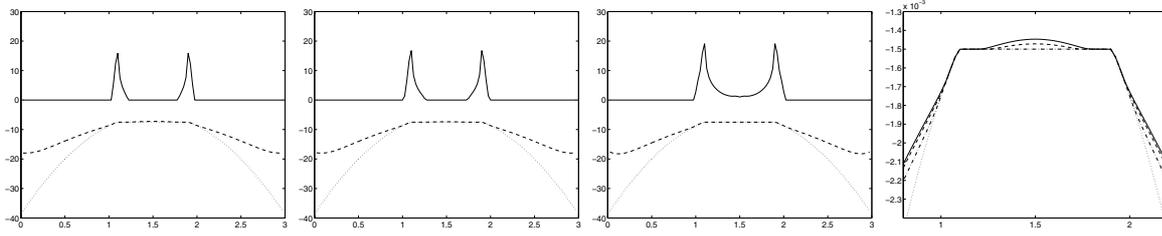


FIGURE 8. Example 2: rigid foundation, normal displacement $\tau_N \mathbf{y}_\gamma$ (both multiplied by 5×10^3) and corresponding multiplier λ_γ for various given friction. From left to right: pure contact problem, $g \equiv 1, 5$ and magnification of the contact zone, $\gamma_1 = \gamma_2 = 10^8$.

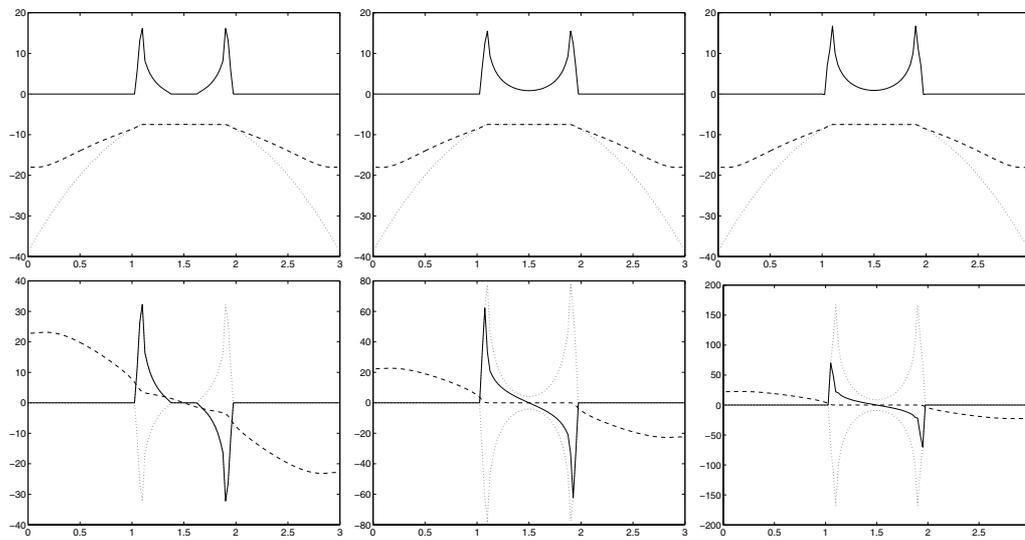


FIGURE 9. Example 2, upper row, left: multiplier λ^* (solid), rigid foundation (multiplied by 5×10^3 , dotted) and normal displacement $\tau_N \mathbf{y}^*$ (multiplied by 5×10^3 , dashed). Second row, left figure: dual variable μ^* (solid) with bounds $\pm \mathfrak{F} \lambda^*$ (dotted) and tangential displacement y^* (multiplied by 5×10^3 , dashed) for $\mathfrak{F} = 2$. Middle column: same as left, but with $\mathfrak{F} = 5$. Right column: same as left, but with $\mathfrak{F} = 10$.

8. CONCLUSIONS AND OUTLOOK

In this paper, primal and dual formulations of 2D contact problems with friction are discussed. The effect of regularization is investigated for both the primal and dual problems and its relation to augmented Lagrangians is discussed. Existence of a solution to a smoothed contact problem with Coulomb friction is shown.

An infinite-dimensional semi-smooth Newton algorithm is applied to the regularized problem with Tresca friction. The method is shown to converge locally superlinear and turns out to be extremely fast and reliable in numerical practice.

Table 1 shows that the number of iterations increases moderately as the parameters γ_1, γ_2 are increased, which motivates a continuation strategy for these parameters. In Table 4, γ_1 and γ_2 are increased heuristically. It would be of importance to derive strategies for an automatic choice and increase of γ_1, γ_2 , *e.g.*, similar to path-following techniques well known in interior point methods.

The algorithms presented in the present work are focused to 2D contact problems with friction. A numerical treatment of frictional contact problems in 3D involves additional difficulties. To be precise, for 3D

problems with friction the dual variable μ in (2.7) is replaced by a vector function $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ that satisfies $\|\boldsymbol{\mu}\| \leq \mathfrak{F}g$ a.e. on Γ_c , i.e., a pointwise *nonlinear* constraint must be considered. The generalization of this paper's methods and analysis to 3D-problems was partly accomplished in [33] and is subject of present research.

Acknowledgements. We wish to thank two anonymous referees for their useful comments which led to this improved version of the paper.

This work was supported in part by the "Fonds zur Förderung der wissenschaftlichen Forschung" under "SFB 03, Optimization and Control".

Appendix A. PROOF OF THEOREM 3.1

Setting $\mathbf{z} := \mathbf{y}_\gamma - \bar{\mathbf{y}}$ in (3.3a) we obtain

$$a(\mathbf{y}_\gamma, \mathbf{y}_\gamma - \bar{\mathbf{y}}) - L(\mathbf{y}_\gamma - \bar{\mathbf{y}}) + (\lambda_\gamma, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} + (\mu_\gamma, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} = 0. \quad (\text{A.1})$$

First we estimate the term $(\lambda_\gamma, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c}$:

$$\begin{aligned} (\lambda_\gamma, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} &= (\lambda_\gamma, \tau_N \mathbf{y}_\gamma - d)_{\Gamma_c} - (\lambda_\gamma, \tau_N \bar{\mathbf{y}} - d)_{\Gamma_c} \\ &\geq \gamma_1^{-1}(\lambda_\gamma, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_\gamma - d))_{\Gamma_c} - \gamma_1^{-1}(\lambda_\gamma, \hat{\lambda})_{\Gamma_c}, \end{aligned}$$

where $(\lambda_\gamma, \tau_N \bar{\mathbf{y}} - d)_{\Gamma_c} \leq 0$ was used. Thus,

$$\begin{aligned} (\lambda_\gamma, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} &\geq \gamma_1^{-1}(\lambda_\gamma, \max(0, \hat{\lambda} + \gamma_1(\tau_N \mathbf{y}_\gamma - d)))_{\Gamma_c} - \gamma_1^{-1}(\lambda_\gamma, \hat{\lambda})_{\Gamma_c} \\ &= \gamma_1^{-1} \|\lambda_\gamma\|_{\Gamma_c}^2 - \gamma_1^{-1}(\lambda_\gamma, \hat{\lambda})_{\Gamma_c} \end{aligned} \quad (\text{A.2})$$

$$= \frac{1}{2\gamma_1} \|\lambda_\gamma - \hat{\lambda}\|_{\Gamma_c}^2 + \frac{1}{2\gamma_1} \|\lambda_\gamma\|_{\Gamma_c}^2 - \frac{1}{2\gamma_1} \|\hat{\lambda}\|_{\Gamma_c}^2 \geq -\frac{1}{2\gamma_1} \|\hat{\lambda}\|_{\Gamma_c}^2. \quad (\text{A.3})$$

For the term $(\mu_\gamma, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c}$ we obtain

$$(\mu_\gamma, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} \leq (\mathfrak{F}g, |\tau_T(\bar{\mathbf{y}} - \mathbf{y}_\gamma)|)_{\Gamma_c} \leq c_1 \|\mathfrak{F}g\|_{\Gamma_c} \|\bar{\mathbf{y}} - \mathbf{y}_\gamma\|_{\mathbf{Y}}, \quad (\text{A.4})$$

with some $c_1 > 0$. Equations (A.1) and (A.2) imply that

$$a(\mathbf{y}_\gamma, \mathbf{y}_\gamma) + \frac{1}{\gamma_1} \|\lambda_\gamma\|_{\Gamma_c}^2 \leq a(\mathbf{y}_\gamma, \bar{\mathbf{y}}) + \frac{1}{\gamma_1}(\lambda_\gamma, \hat{\lambda})_{\Gamma_c} + L(\mathbf{y}_\gamma - \bar{\mathbf{y}}) - (\mu_\gamma, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c}. \quad (\text{A.5})$$

Using (A.4), the coercivity (with constant $c > 0$) and the continuity of $a(\cdot, \cdot)$ in (A.5) shows that

$$c \|\mathbf{y}_\gamma\|_{\mathbf{Y}} + \frac{1}{\gamma_1} \|\lambda_\gamma\|_{\Gamma_c} \quad (\text{A.6})$$

is uniformly bounded with respect to $\gamma_1 \geq 1$. Hence, \mathbf{y}_γ is bounded in \mathbf{Y} and λ_γ in $H^{-1/2}(\Gamma_c)$ using (3.3a). Consequently, there exist $(\tilde{\mathbf{y}}, \tilde{\lambda}) \in \mathbf{Y} \times H^{-1/2}(\Gamma_c)$ and a sequence γ_k with $\lim_{k \rightarrow \infty} \gamma_k = \infty$ such that

$$\mathbf{y}_{\gamma_k} \rightharpoonup \tilde{\mathbf{y}} \text{ weakly in } \mathbf{Y} \quad \text{and} \quad \lambda_{\gamma_k} \rightharpoonup \tilde{\lambda} \text{ weakly in } H^{-\frac{1}{2}}(\Gamma_c). \quad (\text{A.7})$$

Since $|\mu_\gamma| \leq \mathfrak{F}g$ a.e. on Γ_c for all $\gamma_1, \gamma_2 > 0$, there exists $\tilde{\mu}$ and a subsequence γ_{k_l} of γ_k , such that $\mu_{\gamma_{k_l}} \rightharpoonup \tilde{\mu}$ weakly in $L^2(\Gamma_c)$. Since the set $\{\nu \in L^2(\Gamma_c) : |\nu| \leq \mathfrak{F}g \text{ a.e.}\}$ is convex, also the weak limit $\tilde{\mu}$ satisfies $|\tilde{\mu}| \leq \mathfrak{F}g$ almost everywhere. In the sequel we dismiss the subscript k_l with γ_{k_l} . Due to the definition of λ_γ ,

$$\frac{1}{\gamma_1} \|\lambda_\gamma\|_{\Gamma_c}^2 = \gamma_1 \|\max(0, \frac{1}{\gamma_1} \hat{\lambda} + \tau_N \mathbf{y}_\gamma - d)\|_{\Gamma_c}^2.$$

As (A.6) is uniformly bounded with respect to γ_1 , the above inequality implies that

$$\| \max(0, \frac{1}{\gamma_1} \hat{\lambda} + \tau_N \mathbf{y}_\gamma - d) \|_{\Gamma_c}^2 \rightarrow 0 \text{ as } \gamma_1 \rightarrow \infty. \tag{A.8}$$

Since $H^{1/2}(\Gamma_c)$ embeds compactly into $L^2(\Gamma_c)$, (A.7) shows that $\tau_N \mathbf{y}_\gamma \rightarrow \tau_N \tilde{\mathbf{y}}$ almost everywhere on Γ_c . Thus, (A.8) implies that $\tau_N \tilde{\mathbf{y}} - d \leq 0$ almost everywhere on Γ_c , *i.e.* $\tilde{\mathbf{y}} \in \mathbf{d} + \mathbf{K}$.

Subtracting (3.3a) for $(\mathbf{y}_\gamma, \lambda_\gamma, \mu_\gamma)$ from (2.3a) for $(\bar{\mathbf{y}}, \bar{\lambda}, \bar{\mu})$ and setting $\mathbf{z} := \mathbf{y}_\gamma - \bar{\mathbf{y}}$ yields

$$a(\mathbf{y}_\gamma - \bar{\mathbf{y}}, \mathbf{y}_\gamma - \bar{\mathbf{y}}) = -\langle \lambda_\gamma - \bar{\lambda}, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}) \rangle_{\Gamma_c} - (\mu_\gamma - \bar{\mu}, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c}. \tag{A.9}$$

For the second term on the right hand side of (A.9) one deduces

$$(\mu_\gamma - \bar{\mu}, \tau_T \bar{\mathbf{y}})_{\Gamma_c} = (\mu_\gamma, \tau_T \bar{\mathbf{y}})_{\Gamma_c} - (\mathfrak{F}g, |\tau_T \bar{\mathbf{y}}|)_{\Gamma_c} \leq (|\mu_\gamma| - \mathfrak{F}g, |\tau_T \bar{\mathbf{y}}|)_{\Gamma_c} \leq 0.$$

In order to get an estimation for the term $(\bar{\mu} - \mu_\gamma, \tau_T \mathbf{y}_\gamma)$ we distinguish two cases. First, we consider sets where

$$\gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu} \geq \mathfrak{F}g \text{ and thus } \mu_\gamma = \mathfrak{F}g.$$

Then, we have the pointwise estimates

$$(\bar{\mu} - \mu_\gamma) \tau_T \mathbf{y}_\gamma = \frac{1}{\gamma_2} (\bar{\mu} - \mathfrak{F}g) ((\gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu}) - \hat{\mu}) \leq \frac{1}{\gamma_2} \{ \mathfrak{F}g |\hat{\mu}| + |\bar{\mu}| |\hat{\mu}| \}.$$

The same estimate holds on sets where $\gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu} \leq -\mathfrak{F}g$ and $\mu_\gamma = -\mathfrak{F}g$ as well. Let us turn to sets where $|\gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu}| < \mathfrak{F}g$ and $\mu_\gamma = \gamma_2 \tau_T \mathbf{y}_\gamma + \hat{\mu}$. Then, pointwise almost everywhere

$$(\bar{\mu} - \mu_\gamma) \tau_T \mathbf{y}_\gamma = (\bar{\mu} - \gamma_2 \tau_T \mathbf{y}_\gamma - \hat{\mu}) \tau_T \mathbf{y}_\gamma \leq -\gamma_2 |\tau_T \mathbf{y}_\gamma|^2 + \frac{1}{\gamma_2} |\bar{\mu} - \hat{\mu}| (\mathfrak{F}g + |\hat{\mu}|).$$

Combining the above estimates shows that

$$-(\mu_\gamma - \bar{\mu}, \tau_T(\mathbf{y}_\gamma - \bar{\mathbf{y}}))_{\Gamma_c} \leq \frac{1}{\gamma_2} K(\hat{\mu}, \bar{\mu}), \tag{A.10}$$

where $K(\hat{\mu}, \bar{\mu})$ is independent of γ_1, γ_2 . Using (A.3), (A.9), (A.10) and the coercivity of $a(\cdot, \cdot)$ imply that

$$\begin{aligned} 0 &\leq \limsup_{\gamma_1, \gamma_2 \rightarrow \infty} c \| \mathbf{y}_\gamma - \bar{\mathbf{y}} \|_{\mathbf{Y}}^2 \leq \lim_{\gamma_1, \gamma_2 \rightarrow \infty} \{ \langle \bar{\lambda}, \tau_N(\mathbf{y}_\gamma - \bar{\mathbf{y}}) \rangle + \frac{1}{\gamma_2} K(\hat{\mu}, \bar{\mu}) + \frac{1}{\gamma_1} \| \hat{\lambda} \|_{\Gamma_c}^2 \} \\ &= \lim_{\gamma_1, \gamma_2 \rightarrow \infty} \{ \langle \bar{\lambda}, \tau_N \tilde{\mathbf{y}} - d \rangle - \langle \bar{\lambda}, \tau_N \bar{\mathbf{y}} - d \rangle \} = \lim_{\gamma_1, \gamma_2 \rightarrow \infty} \langle \bar{\lambda}, \tau_N \tilde{\mathbf{y}} - d \rangle \leq 0, \end{aligned}$$

where $\tau_N \tilde{\mathbf{y}} - d \leq 0$ on Γ_c was used. From the above estimate it follows that $\mathbf{y}_\gamma \rightarrow \bar{\mathbf{y}}$ strongly in \mathbf{Y} and thus $\tilde{\mathbf{y}} = \bar{\mathbf{y}}$. Passing to the limit in (3.3a) yields that $\tilde{\lambda} = \bar{\lambda}$ and $\tilde{\mu} = \bar{\mu}$. Due to the uniqueness of the solution variables $(\bar{\mathbf{y}}, \bar{\lambda}, \bar{\mu})$ this implies that the whole family $\{(\mathbf{y}_\gamma, \lambda_\gamma, \mu_\gamma)\}$ converges as stated in the theorem.

Appendix B. PROOF OF THEOREM 5.1

Let us denote $\delta_{\mathbf{y}}^l := \mathbf{y}^l - \bar{\mathbf{y}} \in \mathbf{Y}$, $\delta_{\lambda}^l := \lambda^l - \bar{\lambda} \in L^2(\Gamma_c)$ and $\delta_{\mu}^l := \mu^l - \bar{\mu} \in L^2(\Gamma_c)$, where $\bar{\mathbf{y}}$ and $(\bar{\lambda}, \bar{\mu})$ denote the solution variables of (\mathcal{P}) and (\mathcal{P}^*) , respectively. Since $(\bar{\mathbf{y}}, \bar{\lambda}, \bar{\mu})$ satisfies (2.3a), we have for $l \geq 1$ that

$$a(\bar{\mathbf{y}}, \delta_{\mathbf{y}}^{l+1}) - L(\delta_{\mathbf{y}}^{l+1}) + (\bar{\lambda}, \tau_N \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} + (\bar{\mu}, \tau_T \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} = 0, \tag{B.1}$$

and from Step 2 in (ALM) we obtain

$$a(\mathbf{y}^{l+1}, \delta_{\mathbf{y}}^{l+1}) - L(\delta_{\mathbf{y}}^{l+1}) + (\lambda^{l+1}, \tau_N \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} + (\mu^{l+1}, \tau_T \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} = 0. \tag{B.2}$$

Subtracting (B.1) from (B.2) results in

$$\begin{aligned} 0 &= a(\mathbf{y}^{l+1} - \bar{\mathbf{y}}, \delta_{\mathbf{y}}^{l+1}) + (\lambda^{l+1} - \bar{\lambda}, \tau_N \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} + (\mu^{l+1} - \bar{\mu}, \tau_T \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} \\ &= a(\delta_{\mathbf{y}}^{l+1}, \delta_{\mathbf{y}}^{l+1}) + (\delta_{\lambda}^{l+1}, \tau_N \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} + (\delta_{\mu}^{l+1}, \tau_T \delta_{\mathbf{y}}^{l+1})_{\Gamma_c}. \end{aligned} \tag{B.3}$$

Due to Step 2 of (ALM), (2.3b) and (2.3c) and the assumption that $\bar{\lambda} \in L^2(\Gamma_c)$ we have

$$\lambda^{l+1} = P_1(\lambda^l + \gamma_1^{l+1}(\tau_N \mathbf{y}^{l+1} - d)), \quad \bar{\lambda} = P_1(\bar{\lambda} + \gamma_1^{l+1}(\tau_N \bar{\mathbf{y}} - d)), \tag{B.4}$$

$$\mu^{l+1} = P_2(\mu^l + \gamma_2^{l+1} \tau_T \mathbf{y}^{l+1}) \quad \text{and} \quad \bar{\mu} = P_2(\bar{\mu} + \gamma_2^{l+1} \tau_T \bar{\mathbf{y}}), \tag{B.5}$$

where $P_1 : L^2(\Gamma_c) \rightarrow L^2(\Gamma_c)$ denotes the pointwise projection onto the convex set $K_1 = \{\xi \in L^2(\Gamma_c) : \xi \geq 0 \text{ a.e.}\}$, and $P_2 : L^2(\Gamma_c) \rightarrow L^2(\Gamma_c)$ the projection onto $K_2 = \{\nu \in L^2(\Gamma_c) : |\nu| \leq \mathfrak{F}g \text{ a.e.}\}$. From the properties of projections onto convex sets, one gets that

$$\begin{aligned} &(\lambda^{l+1} - \bar{\lambda}, (\lambda^l + \gamma_1^{l+1}(\tau_N \mathbf{y}^{l+1} - d)) - \lambda^{l+1}) - (\bar{\lambda} + \gamma_1^{l+1}(\tau_N \bar{\mathbf{y}} - d) - \bar{\lambda})_{\Gamma_c} \geq 0, \\ &(\mu^{l+1} - \bar{\mu}, (\mu^l + \gamma_2^{l+1} \tau_T \mathbf{y}^{l+1} - \mu^{l+1}) - (\bar{\mu} + \gamma_2^{l+1} \tau_T \bar{\mathbf{y}} - \bar{\mu}))_{\Gamma_c} \geq 0. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} &(\delta_{\lambda}^{l+1}, \tau_N \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} + (\delta_{\mu}^{l+1}, \tau_T \delta_{\mathbf{y}}^{l+1})_{\Gamma_c} = (\gamma_1^{l+1})^{-1} (\lambda^{l+1} - \bar{\lambda}, (\lambda^l + \gamma_1^{l+1}(\tau_N \mathbf{y}^{l+1} - d)) - (\bar{\lambda} + \gamma_1^{l+1}(\tau_N \bar{\mathbf{y}} - d)))_{\Gamma_c} \\ &\quad + (\gamma_2^{l+1})^{-1} (\mu^{l+1} - \bar{\mu}, (\mu^l + \gamma_2^{l+1} \tau_T \mathbf{y}^{l+1}) - (\bar{\mu} + \gamma_2^{l+1} \tau_T \bar{\mathbf{y}}))_{\Gamma_c} \\ &\quad - (\gamma_1^{l+1})^{-1} (\lambda^{l+1} - \bar{\lambda}, \lambda^l - \bar{\lambda})_{\Gamma_c} - (\gamma_2^{l+1})^{-1} (\mu^{l+1} - \bar{\mu}, \mu^l - \bar{\mu})_{\Gamma_c} \\ &\geq (\gamma_1^{l+1})^{-1} \|\lambda^{l+1} - \bar{\lambda}\|_{\Gamma_c}^2 - (\gamma_1^{l+1})^{-1} (\lambda^{l+1} - \bar{\lambda}, \lambda^l - \bar{\lambda})_{\Gamma_c} \\ &\quad + (\gamma_2^{l+1})^{-1} \|\mu^{l+1} - \bar{\mu}\|_{\Gamma_c}^2 - (\gamma_2^{l+1})^{-1} (\mu^{l+1} - \bar{\mu}, \mu^l - \bar{\mu})_{\Gamma_c} \\ &\geq \frac{1}{2\gamma_1^{l+1}} \|\delta_{\lambda}^{l+1}\|_{\Gamma_c}^2 - \frac{1}{2\gamma_1^{l+1}} \|\delta_{\lambda}^l\|_{\Gamma_c}^2 + \frac{1}{2\gamma_2^{l+1}} \|\delta_{\mu}^{l+1}\|_{\Gamma_c}^2 - \frac{1}{2\gamma_2^{l+1}} \|\delta_{\mu}^l\|_{\Gamma_c}^2. \end{aligned} \tag{B.6}$$

Utilizing (B.3), (B.6) and $\gamma_1^l \leq \gamma_1^{l+1}$, $\gamma_2^l \leq \gamma_2^{l+1}$ yields that

$$\frac{1}{2\gamma_1^{l+1}} \|\delta_{\lambda}^{l+1}\|_{\Gamma_c}^2 + \frac{1}{2\gamma_2^{l+1}} \|\delta_{\mu}^{l+1}\|_{\Gamma_c}^2 \leq \frac{1}{2\gamma_1^l} \|\delta_{\lambda}^l\|_{\Gamma_c}^2 + \frac{1}{2\gamma_2^l} \|\delta_{\mu}^l\|_{\Gamma_c}^2 - a(\delta_{\mathbf{y}}^{l+1}, \delta_{\mathbf{y}}^{l+1}).$$

This shows that $a(\delta_{\mathbf{y}}^l, \delta_{\mathbf{y}}^l) \rightarrow 0$ as $\gamma_1, \gamma_2 \rightarrow \infty$, and thus $\mathbf{y}^l \rightarrow \bar{\mathbf{y}}$ strongly in \mathbf{Y} . From Step 2 of (ALM) it follows that $\lambda^l \rightarrow \bar{\lambda}$ and $\mu^l \rightarrow \bar{\mu}$ weakly in $L^2(\Gamma_c)$, as desired.

REFERENCES

- [1] P. Alart and A. Curnier, A mixed formulation for frictional contact problems prone to Newton like solution methods. *Comput Methods Appl. Mech. Engrg.* **92** (1991) 353–375.
- [2] J. Albery, C. Carstensen, S.A. Funken and R. Klose, Matlab implementation of the finite element method in elasticity. *Computing* **69** (2002) 239–263.
- [3] A. Bensoussan and J. Frehse, *Regularity results for nonlinear elliptic systems and applications*, Springer-Verlag, Berlin. *Appl. Math. Sci.* **151** (2002).
- [4] M. Bergounioux, M. Haddou, M. Hintermüller and K. Kunisch, A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM J. Optim.* **11** (2000) 495–521.
- [5] X. Chen, Z. Nashed and L. Qi, Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM J. Numer. Anal.* **38** (2000) 1200–1216.
- [6] P.W. Christensen and J.S. Pang, Frictional contact algorithms based on semismooth Newton methods, in *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods*, Kluwer Acad. Publ., Dordrecht. *Appl. Optim.* **22** (1999) 81–116.
- [7] P.W. Christensen, A. Klarbring, J.S. Pang and N. Strömberg, Formulation and comparison of algorithms for frictional contact problems. *Internat. J. Numer. Methods Engrg.* **42** (1998) 145–173.
- [8] Z. Dostál, J. Haslinger and R. Kučera, Implementation of the fixed point method in contact problems with Coulomb friction based on a dual splitting type technique. *J. Comput. Appl. Math.* **140** (2002) 245–256.
- [9] C. Eck and J. Jarušek, Existence results for the static contact problem with Coulomb friction. *Math. Models Methods Appl. Sci.* **8** (1998) 445–468.
- [10] I. Ekeland and R. Témam, *Convex analysis and variational problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. *Classics in Applied Mathematics* **28** (1999).
- [11] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, second edition. *Grundlehren der Mathematischen Wissenschaften* **224** (1983).
- [12] R. Glowinski, *Numerical methods for nonlinear variational problems*. Springer Series in Computational Physics. Springer-Verlag, New York (1984).
- [13] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman (Advanced Publishing Program), Boston, MA. *Monographs Stud. Math.* **24** (1985).
- [14] W. Han and M. Sofonea, *Quasistatic contact problems in viscoelasticity and viscoplasticity*, American Mathematical Society, Providence, RI. *AMS/IP Studies in Advanced Mathematics* **30** (2002).
- [15] J. Haslinger, Approximation of the Signorini problem with friction, obeying the Coulomb law. *Math. Methods Appl. Sci.* **5** (1983) 422–437.
- [16] J. Haslinger, Z. Dostál and R. Kučera, On a splitting type algorithm for the numerical realization of contact problems with Coulomb friction. *Comput. Methods Appl. Mech. Engrg.* **191** (2002) 2261–2281.
- [17] M. Hintermüller and M. Ulbrich, A mesh-independence result for semismooth Newton methods. *Math. Prog., Ser. B* **101** (2004) 151–184.
- [18] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2003) 865–888.
- [19] M. Hintermüller, V. Kovtunenkov and K. Kunisch, Semismooth Newton methods for a class of unilaterally constrained variational inequalities. *Adv. Math. Sci. Appl.* **14** (2004) 513–535.
- [20] I. Hlaváček, J. Haslinger, J. Nečas and J. Lovíšek, *Solution of Variational Inequalities in Mechanics*. Springer, New York. *Appl. Math. Sci.* **66** (1988).
- [21] S. Hübner and B. Wohlmuth, A primal-dual active strategy for non-linear multibody contact problems. *Comput. Methods Appl. Mech. Engrg.* **194** (2005) 3147–3166.
- [22] K. Ito and K. Kunisch, Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces. *Nonlinear Anal.* **41** (2000) 591–616.
- [23] K. Ito and K. Kunisch, Semi-smooth Newton methods for variational inequalities of the first kind. *ESAIM: M2AN* **37** (2003) 41–62.
- [24] N. Kikuchi and J.T. Oden, *Contact problems in elasticity: a study of variational inequalities and finite element methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. *SIAM Stud. Appl. Math.* **8** (1988).
- [25] A. Klarbring, Mathematical programming and augmented Lagrangian methods for frictional contact problems, A. Curnier, Ed. *Proc. Contact Mechanics Int. Symp.* (1992).
- [26] R. Krause, *Monotone Multigrid Methods for Signorini's Problem with Friction*. Ph.D. Thesis, FU Berlin (2001).
- [27] P. Laborde and Y. Renard, *Fixed point strategies for elastostatic frictional contact problems*. Rapport Interne 03-27, MIP Laboratory, Université Paul Sabatier, Toulouse (2003).
- [28] A.Y.T. Leung, Guoqing Chen and Wanji Chen, Smoothing Newton method for solving two- and three-dimensional frictional contact problems. *Internat. J. Numer. Methods Engrg.* **41** (1998) 1001–1027.

- [29] C. Licht, E. Pratt and M. Raous, Remarks on a numerical method for unilateral contact including friction, in *Unilateral problems in structural analysis, IV (Capri, 1989)*, Birkhäuser, Basel. *Internat. Ser. Numer. Math.* **101** (1991) 129–144.
- [30] J. Nečas, J. Jarušek and J. Haslinger, On the solution of the variational inequality to the Signorini problem with small friction. *Boll. Unione Math. Ital.* **5** (1980) 796–811.
- [31] C.A. Radoslovescu and M. Cocu, Internal approximation of quasi-variational inequalities. *Numer. Math.* **59** (1991) 385–398.
- [32] M. Raous, Quasistatic Signorini problem with Coulomb friction and coupling to adhesion, in *New developments in contact problems*, P. Wriggers and Panagiotopoulos, Eds., Springer Verlag. *CISM Courses and Lectures* **384** (1999) 101–178.
- [33] G. Stadler, *Infinite-Dimensional Semi-Smooth Newton and Augmented Lagrangian Methods for Friction and Contact Problems in Elasticity*. Ph.D. Thesis, University of Graz (2004).
- [34] G. Stadler, Semismooth Newton and augmented Lagrangian methods for a simplified friction problem. *SIAM J. Optim.* **15** (2004) 39–62.
- [35] M. Ulbrich, Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim.* **13** (2003) 805–842.