Charles University in Prague

Faculty of Mathematics and Physics

DOCTORAL THESIS

Róbert Pathó

Shape optimization in contact problems with friction

Department of Numerical Mathematics

Supervisor of the doctoral thesis:	prof. RNDr. Jaroslav Haslinger, DrSc.
Study programme:	Mathematics
Specialization:	Scientific and Technical Calculations

Prague 2014

Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE

Róbert Pathó

Tvarová optimalizace v kontaktních úlohách se třením

Katedra numerické matematiky

Vedoucí disertační práce: prof. RNDr. Jaroslav Haslinger, DrSc.

Studijní program: Matematika

Studijní obor: Vědecko-technické výpočty

Praha 2014

Hereby I would like to express my thanks to prof. RNDr. Jaroslav Haslinger, DrSc. for his patient guidance throughout my studies and doc. Ing. Jiří V. Outrata, DrSc. for introducing me to the exciting world of variational analysis. It was a delight to collaborate with Ing. Petr Beremlijski, Ph.D. on the numerical part, for which I am deeply indebted. Next, I would also like to thank prof. Marko M. Mäkelä for providing recent information on nonsmooth optimization algorithms and for his proximal bundle solver. Finally, the greatest appretiation goes to my family and girlfriend, without whose love and support this thesis would not have been possible. Financial support from project no. P201/12/0671 of the GAČR and project no. 719912 of GAUK is gratefully acknowledged.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

Prague, 25th June 2014 Róbert Pathó

Název práce: Tvarová optimalizace v kontaktních úlohách se třením

Autor: Róbert Pathó

Katedra: Katedra numerické matematiky

Vedoucí disertační práce: prof. RNDr. Jaroslav Haslinger, DrSc.

Abstrakt: Cílem práce je nalézt optimální tvar elastického tělesa, která je v (statickém) kontaktu s dokonale tuhou překážkou. Na kontaktní části hranice uvažujeme dva modely tření: Trescův a Coulombův zákon tření, kde ovšem koeficient tření může záviset na velikosti neznámého tečného posunutí. V diskretizované úloze je kontaktní hranice popsána konečným počtem parametrů, tzv. návrhový vektor, a stavovou úlohu tvoří (v obou uvažovaných případech) konečně-dimenzionální implicitni variační nerovnice druhého druhu, parametrizována tímto návrhovým vektorem. V práci ukážeme, že v jisté přípustné množině optimální tvar existuje pro libovolnou "rozumnou"cenovou funkci, a navrhneme vhodnou metodu pro jeho výpočet. Ta je založena na kombinaci implicitního programování a analýze citlivosti, která umožňuje použití efektivních minimalizačních algoritmů. Aplikovatelnost zvoleného přístupu je demonstrována na několika konkrétních příkladech.

Klíčová slova: tvarová optimalizace; kontaktní úloha; koeficient tření, který závisí na řešení; analýza citlivosti; Mordukhovichův kalkulus

Title: Shape optimization in contact problems with friction

Author: Róbert Pathó

Department: Department of Numerical Mathematics

Supervisor: prof. RNDr. Jaroslav Haslinger, DrSc.

Abstract: The aim of the present thesis is to find an optimal shape of an elastic body that is in (static) contact with a rigid obstacle. On the contact boundary we assume two models of friction: the Tresca and Coulomb laws of friction, in which the coefficient of friction may depend on the unknown tangential displacement. In the discretized problem the contact boundary is described by a finite number of parameters, the so-called design vector, and the state problem is represented by a finite-dimensional implicit variational inequality of the second kind, that is parametrized by the design vector. We show that, given a suitable admissible set, an optimal shape exists for every "reasonable"cost functional, and propose an algorithm for its computation. To this end we combine the implicit programming approach with sensitivity analysis, facilitating the use of effective minimization methods. The applicability of the proposed method is demonstrated on several numerical examples.

Keywords: shape optimization; contact problem; solution-dependent coefficient of friction; sensitivity analysis; Mordukhovich calculus

Contents

List of Figures

Notation

Sets

$\emptyset$	empty set
$\mathbb{N}$	set of positive integers $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	set of nonnegative integers $\{0, 1, 2, \dots\}$
$\mathbb{R}$	set of real numbers $(-\infty, +\infty)$
$\mathbb{R}_+$	set of nonnegative real numbers $[0, +\infty)$
$\mathbb{R}_-$	set of nonpositive real numbers $(-\infty, 0]$
(a, b)	open interval in $\mathbb{R}$
$[a, b]$	closed interval in $\mathbb{R}$
$\mathbb{B}_r(\boldsymbol{x})$	closed ball of radius $r > 0$ and center $\boldsymbol{x} \in \mathbb{R}^n$
$\mathbb{B}_r$	$= \mathbb{B}_r(\mathbf{0})$
$X \times Y$	Cartesian product of the sets X and Y
X^n	$\underbrace{X \times X \times \cdots \times X}_{n \text{ times}}$
$\overline{A}$	closure of $A \subset X$ (in the topology of X)
∂A	topological boundary of $A \subset X$
$\operatorname{relint} A$	relative interior of $A \subset X$
$\operatorname{conv} A$	convex hull of A
$\{x^i\}_{i \in I}$, $\{x^{(i)}\}_{i \in I}$	sequence of elements ($I \subset \mathbb{N}$)

Functions and mappings

$f : X \to \mathbb{R}$	real-valued function from X into $\mathbb{R}$
$\boldsymbol{f} : X \to \mathbb{R}^n$	vector-valued mapping from X into $\mathbb{R}^n$
$F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$	multifunction (or set-valued mapping); a mapping from $\mathbb{R}^n$ into subsets of $\mathbb{R}^m$
$x \mapsto f(x)$	function (mapping) f
$C^k(\Omega)$	space of functions having continuous derivatives on Ω up to order $k \in \mathbb{N}_0$
$C^{0,1}(\Omega)$	space of Lipschitz functions on Ω
$L^p(\Omega)$	Lebesgue integrable functions of order $p \geq 1$ on Ω
$L^\infty(\Omega)$	essentially bounded measurable functions on Ω
$H^1(\Omega)$	Sobolev space $W^{1,2}(\Omega)$ of functions belonging together with their distributional derivatives into $L^2(\Omega)$
$\operatorname{div} \boldsymbol{f}$	divergence operator
∇f	gradient of a real-valued function
$\nabla \boldsymbol{f}$	Jacobian matrix of a vector-valued function
$\overline{\partial} f$	Clarke's subgradient of a real-valued f
$\overline{\partial} \boldsymbol{f}$	Clarke's generalized Jacobian of vector-valued $\boldsymbol{f}$
∂f	limiting subdifferential of the real-valued function f
D^*Q	limiting coderivative of the multifunction Q
$\operatorname{dist}(\boldsymbol{x}, S)$	distance of $\boldsymbol{x} \in \mathbb{R}^n$ from the set $S \subset \mathbb{R}^n$

Linear algebra

$\mathbb{R}^n$	Euclidean space of dimension n
$\boldsymbol{x} \in \mathbb{R}^n$	column vector $\boldsymbol{x} = (x_1, \ldots, x_n)^T$
x_i	ith component of $\boldsymbol{x} \in \mathbb{R}^n$
$\mathbb{R}^{n \times m}$	space of matrices of type $n \times m$
$\mathbb{E}$	unit matrix
$\mathbb{A}^T$	transposed matrix
$\mathbb{A}^{-1}$	inverse matrix
$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_n = \boldsymbol{x} \cdot \boldsymbol{y}$	$= \sum_{i=1}^n x_i y_i$; Euclidean scalar product on $\mathbb{R}^n$
$\Vert\boldsymbol{x}\Vert_n$	$= (\boldsymbol{x} \cdot \boldsymbol{x})^{1/2}$; Euclidean norm on $\mathbb{R}^n$
$\Vert\boldsymbol{x}\Vert_\infty$	$= \max\{\lvert x_i\rvert \mid i = 1, \ldots, n\}$; max-norm on $\mathbb{R}^n$
$\boldsymbol{x} \bullet \boldsymbol{y}$	$= (x_1 y_1, \ldots, x_n y_n)^T$; componentwise product on $\mathbb{R}^n$
$\boldsymbol{x} \div \boldsymbol{y}$	$= (x_1/y_1, \ldots, x_n/y_n)^T$; componentwise division on $\mathbb{R}^n$
$\lvert\boldsymbol{x}\rvert$	$= (\lvert x_1\rvert, \ldots, \lvert x_n\rvert)^T$; componentwise absolute value
$\boldsymbol{x} \geq \boldsymbol{y}$	componentwise comparison, i.e. $x_i \geq y_i \; \forall i = 1, \ldots, n$;
$\mathbb{A} : \mathbb{B}$	$= \sum_{i=1}^n \sum_{j=1}^m A_{ij} B_{ij}$; scalar product on $\mathbb{R}^{n \times m}$

Abbreviations

AGE	adjoint generalized equation
BT	bundle trust
CP	cutting plane
GE	generalized equation
ImP	implicit programming
MPEC	mathematical program with equilibrium constraints
NLP	nonlinear program
SRC	strong regularity condition

Introduction

There is virtually no area of mechanical engineering where one would not encounter the problem of determining the motion/position of several deformable bodies that are or may become in contact, but cannot penetrate one another. Moreover, in many applications it is simply not possible to neglect the action of friction forces on the contacting surfaces. These frictional effects may be welcomed, e.g. in machine tools, or undesirable, e.g. because they cause wear in the material and thus shorten the lifecycle of the contacting parts. In either case, engineers have always tried to maximize the desired effects just by altering the geometry of the modelled elements—this is the topic of contact shape optimization.

From the mathematical point of view, *shape optimization* is the branch of optimal control theory, where the control variable (also called *design variable* in the context of shape optimization) is connected to the geometry of the problem. A fundamental role in shape optimization problems is played by the *control-to-state mapping* (or *solution map*) S, which assigns to each feasible value of the design variable the set of solutions to the *state problem*. Thus, any shape optimization problem can be written in the following general form:

$$\left.\begin{array}{ll} \text{minimize} & J(\alpha, y), \\ \text{subj. to} & y \in S(\alpha), \\ & \alpha \in U_{ad}, \end{array}\right\} \tag{1}$$

where the real-valued function J is called the *cost functional*, U_{ad} signifies an admissible set of design variables α and S usually represents an equilibrium problem. Typically, the state variable y is sought in a function space $V(\alpha)$, where α determines the domain of definition of y. After suitable discretization, (1) turns into a (finite-dimensional) *mathematical program with equilibrium constraints* (MPEC), where $\alpha \in U_{ad} \subset \mathbb{R}^m$, $y \in \mathbb{R}^n$ and $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$. If the state problems happen to be uniquely solvable for each α, i.e. S is single-valued, we can substitute $y = S(\alpha)$ and solve the MPEC as a standard nonlinear optimization problem. This is called the *implicit programming* (ImP) approach, since the composite cost functional $\mathcal{J} : \alpha \mapsto J(\alpha, S(\alpha))$ involves the implicitly defined control-to-state mapping. Fast minimization algorithms may be applied provided one is able to compute (sub)gradients of $\mathcal{J}$. As it turns out, this is a major problem whenever S is complicated enough.

Due to its importance, shape optimization in contact problems has been subject to research for quite some time—let us mention the monographs [15, 16, 52] and the references therein. For example, in [16] the two-dimensional (2D) Signorini problem is considered without friction and with Tresca friction; the papers [3] and [4] analyse the same problem with Coulomb friction in two and three dimensions, respectively. In the present thesis we aim at generalizing these results to 2D contact problems with Tresca and Coulomb laws of friction, where the coefficient of friction may depend on the magnitude of the tangential displacement. By means of it one can model, e.g., in dynamic contact problems the transition from the static friction coefficient to the dynamic one, or the stick-slip motion during earthquakes—see e.g. [49]. In their weak formulation, contact problems with Tresca friction and a solution-dependent coefficient of friction take the form of an implicit variational inequality of the second kind, similarly as for the local Coulomb law in 2D in [3]; however, in our case it cannot be proved that the control-to-state mapping S is piecewise smooth, unless imposing additional smoothness assumptions on the coefficient of friction. Therefore, when performing sensitivity analysis we follow rather [4] and employ the generalized differential calculus of B. Mordukhovich to derive first order sensitivities of S. Moreover, in contrast to the Coulomb case, the discretized state problem is formulated as a generalized equation with a control-dependent multivalued part, which is a rather uncommon model in the literature. Things get even more complicated as the local Coulomb law of friction is coupled with a solution-dependent friction coefficient; nevertheless, the established approach proved to be successful also in this case.

The thesis consists of four chapters and an appendix. In Chapter 1 we introduce the state problems while keeping the design variable α fixed. We deal with frictional contact problems and their various variational formulations, discretization and solvability with respect to the coefficient of friction.

Chapter 2 deals with shape optimization in contact problems with the Tresca model of friction and a solution-dependent coefficient of friction. After recalling briefly the main results of [43] we move onto the discrete shape optimization problem, prove its solvability and conduct sensitivity analysis based on modern tools from variational analysis.

The structure of Chapter 3, where we investigate shape opimization in contact problems with Coulomb friction and a solution-dependent coefficient of friction, very much resembles that of Chapter 2. However, in this case we treat only the discretized shape optimization problem: Lipschitz continuity of S and solvability of the shape optimization problem is proved. Again, sensitivity analysis represents the core of the chapter, providing for a subgradient of the cost functional in numerical experiments.

Chapter 4 introduces first the tools used for the numerical realization of contact shape optimization problems: the bundle trust minimization algorithm, solution of the state problems and the adjoint equations, then finally several examples are presented.

For the sake of completeness and convenience of the reader, we provide a summary of those basic tools from nonsmooth and variational analysis (Clarke's and Mordukhovich's calculus) in Appendix A that are used extensively throughout Chapters 2 and 3.

Chapter 1

Contact problems with various models of friction

In the following introductory chapter we describe the state problems, parametrized by the geometry of the underlying domain. This will play central role in the subsequent chapters dealing with finding an optimal value of this parameter. We start our exposition with the classical Signorini problem in linearized elasticity (posed originally in [51] and solved in [13, 14], paving the ground for the theory of variational inequalities) combined with the most basic model of friction, the so-called Tresca law. Due to its simplicity, this problem has been thoroughly analyzed and questions concerning its (unique) solvability answered satisfactorily (see e.g. the monographs [11, 22, 24] on unilateral contact problems). Based on the aforementioned problem we introduce and analyze properties of contact problems with "generalized" Tresca and Coulomb laws of friction, where we allow the coefficient of friction $\mathfrak{F}$ to depend on the unknown solution (see also [19, 20] and [30, 31] for the three-dimensional case). In particular, the classical and weak formulations of these frictional contact problems shall be presented, followed by their finite element discretization. Conditions guaranteeing existence and uniqueness of the corresponding solutions will be recalled. Note, that in shape optimization one does not deal with a particular state problem, but rather a family of problems, which differ in their geometry, i.e. the domain of definition of the unknown solution. Therefore, throughout the presentation below a parameter α will occur, that determines the shape of the underlying domain. Special attention will paid to the unique solvability of the discrete state problems with respect to α.

1.1 The Signorini problem with given friction

We start with some basic notions from the theory of linearized elasticity and contact mechanics. Let a planar, elastic body, in its reference configuration, be represented by the domain $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary $\partial\Omega$ (later the geometry of Ω will be further specified). Assume that $\partial\Omega$ is composed of three nonempty, pairwise disjoint and relatively open parts Γ_D, Γ_N, Γ_C so that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N \cup \overline{\Gamma}_C$. The body Ω is subject to volume forces of density $\boldsymbol{F} : \Omega \to \mathbb{R}^2$ and surface tractions of density $\boldsymbol{P} : \Gamma_N \to \mathbb{R}^2$, while Ω is clamped on Γ_D. The classical *Signorini problem* consists in finding a displacement field $\boldsymbol{u} : \overline{\Omega} \to \mathbb{R}^2$ such that the deformed body is in equilibrium with the forces acting upon it, whereas, in addition, one assumes the presence of a *perfectly rigid obstacle* $\Xi \subset \mathbb{R}^2$ (i.e., Ξ does not undergo any deformation). The so-called *contact boundary* Γ_C is the part of $\partial\Omega$, where Ω *may* become in contact with Ξ, but can not penetrate into it (see Figure 1.1). In the classical Signorini problem the contact along Γ_C is assumed to be frictionless, but it will not be our case.

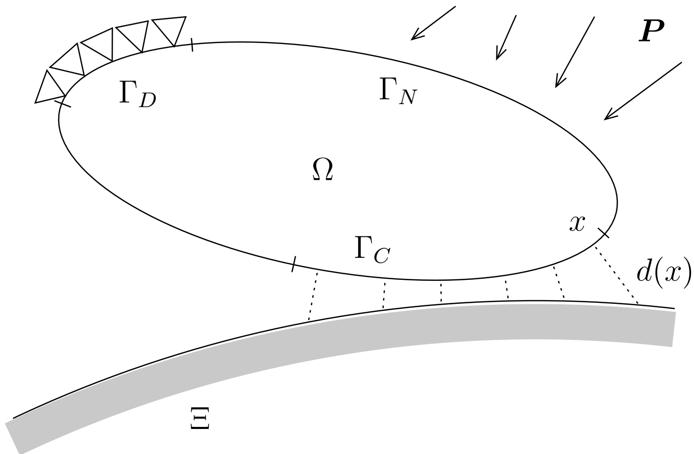


Figure 1.1: 2D Signorini problem.

In order to present the differential equations the unknown displacement field $\boldsymbol{u}$ has to satisfy, we introduce the following notation: $\varepsilon(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T)$ shall denote the linearized strain tensor; the stress tensor $\sigma : \Omega \to \mathbb{R}^{2\times 2}$ will be linked to ε by means of a linear Hooke's law, determined by the fourth-order stiffness tensor $\mathcal{C} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$. Denoting by $\boldsymbol{\nu} : \partial\Omega \to \mathbb{R}^2$ the unit outward normal to $\partial\Omega$, one defines the *normal component* of the displacement along $\partial\Omega$ as $u_n := (\boldsymbol{u}|_{\partial\Omega}) \cdot \boldsymbol{\nu}$ and the *stress vector* as $\boldsymbol{T} := \sigma\boldsymbol{\nu}$. The *normal stress* is defined as the normal component of the stress vector, i.e. $T_n := \sigma\boldsymbol{\nu}\cdot\boldsymbol{\nu}$ and the *tangential stress* as $\boldsymbol{T}_t := \boldsymbol{T} - T_n\boldsymbol{\nu}$.

By the *classical solution* to the Signorini problem with Tresca friction we mean any displacement field $\boldsymbol{u} : \overline{\Omega} \to \mathbb{R}^2$ satisfying the following system of differential equations and boundary conditions (abbreviated as b.c. below):

(equilibrium equation)

$$\operatorname{div}\sigma + \boldsymbol{F} = \boldsymbol{0} \quad \text{in } \Omega, \tag{1.1}$$

(Hooke's law)

$$\sigma = \mathcal{C}\varepsilon(\boldsymbol{u}) \quad \text{in } \Omega, \tag{1.2}$$

(Dirichlet b.c.)

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma_D, \tag{1.3}$$

(Neumann b.c.)

$$\boldsymbol{T} = \boldsymbol{P} \quad \text{on } \Gamma_N, \tag{1.4}$$

(unilateral b.c.)

$$u_n \leq d, \quad T_n \leq 0, \quad T_n(u_n - d) = 0 \quad \text{on} \quad \Gamma_C, \tag{1.5}$$

(frictional b.c.)

$$\|\boldsymbol{T}_t\| \leq \mathfrak{F}g, \quad \boldsymbol{u}_t \neq \boldsymbol{0} \Rightarrow \boldsymbol{T}_t = -\mathfrak{F}g\frac{\boldsymbol{u}_t}{\|\boldsymbol{u}_t\|} \quad \text{on} \quad \Gamma_C. \tag{1.6}$$

The function $d : \Gamma_C \to \mathbb{R}_+$ appearing in (1.5) is called the *gap* (or distance) *function* and the first inequality $(1.5)_1$ models the fact that the gap between the deformed body and the rigid obstacle is positive or equal to zero; by $(1.5)_2$ we exclude adhesion (only compression is allowed); finally, the complementarity condition $(1.5)_3$ says that pressure may occur only at points of contact. Here we complement the contact boundary conditions on Γ_C with the simplest model of friction, the *Tresca law*, or the so-called model with *given friction* (1.6). It says that no slip occurs until the shear stress does not attain a certain threshold value, given by the product of the *coefficient of friction* $\mathfrak{F} : \Gamma_C \to \mathbb{R}_+$ and an a priori given function $g : \Gamma_C \to \mathbb{R}_+$ called the *slip bound*. Note that (1.6) is merely a simplification of the physically more relevant and widely used *Coulomb law* of friction, that will be introduced and discussed later in this chapter.

Now we specify the geometrical setting of the contact problem that will be dealt with in the sequel. In particular, we assume that the rigid obstacle is *flat* and the elastic body is represented by a "rectangle" with curved contact zone only (since our goal is to optimize the contact boundary, this does not represent a relevant simplification)–see Figure 1.2. Therefore, by a suitable choice of the coordinate system, $\Xi = \mathbb{R} \times \mathbb{R}_-$ (recall, that Ξ denotes the obstacle) and $\Omega \subset \widehat{\Omega} := (0, a) \times (0, b)$, as shown in Figure 1.2. Further, we assume that Γ_C can be described by one Lipschitz continuous function α, i.e. $\overline{\Gamma}_C = \operatorname{Gr} \alpha$. This parameter

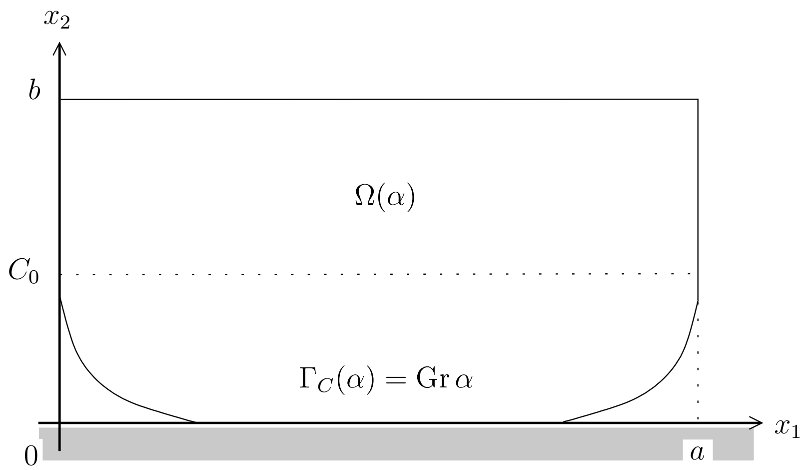


Figure 1.2: Geometry of our contact problem.

α, called the *design variable* in context of optimal shape design, is going to be subject to optimization in the forthcoming chapters. An optimal α will be sought in the admissible set

$$\begin{aligned} \mathcal{U}_{ad} := \Big\{ \alpha \in C^{0,1}([0,a]) \;\Big|\; & 0 \le \alpha \le C_0 \ \text{ in } \ [0,a], \\ & |\alpha'| \le C_1 \ \text{ a.e. in } \ (0,a), \\ & C_{21} \le \int_0^a \alpha(x_1)\, dx_1 \le C_{22} \Big\}. \end{aligned} \tag{1.7}$$

We assume that the positive constants C_0, C_1, C_{21}, C_{22} are given in such a way that $\mathcal{U}_{ad} \neq \emptyset$. Thus the elastic body Ω with $\overline{\Gamma}_C = \operatorname{Gr} \alpha$ becomes

$$\Omega = \Omega(\alpha) := \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 < x_1 < a,\ \alpha(x_1) < x_2 < b\}$$

and the third condition in (1.7) translates to $\widetilde{C}_{21} \leq \operatorname{meas} \Omega(\alpha) \leq \widetilde{C}_{22}$ with $\widetilde{C}_{21} = ab - C_{22}$ and $\widetilde{C}_{22} = ab - C_{21}$. In particular, by setting $C_{21} = C_{22}$ one may enforce that all admissible bodies in

$$\mathcal{O} := \{\Omega(\alpha) \mid \alpha \in \mathcal{U}_{ad}\}$$

have the same volume.

Next, let us reformulate the general contact conditions (1.5) and (1.6), exploiting the special geometry described above. First of all, note that the inequality $(1.5)_1$ represents only an approximation of the nonpenetration condition between two bodies in the framework of small deformations. In general, $(1.5)_1$ does not guarantee that the deformed body stays above the obstacle, e.g. Figure 1.3 depicts an example of Ω penetrating into Ξ. Therefore, we will consider a modified ver-

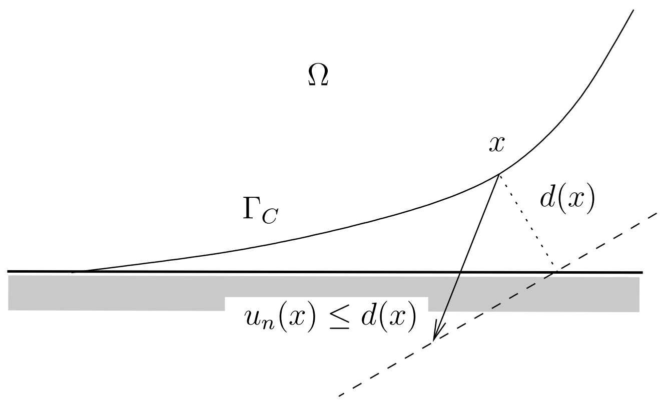


Figure 1.3: Penetration into the obstacle.

sion of the complementarity system (1.5), which ensures nonpenetration *exactly*, along the whole Γ_C:

$$-u_2(\boldsymbol{x}) \leq x_2, \quad T_2(\boldsymbol{x}) \geq 0, \quad T_2(\boldsymbol{x})(u_2(\boldsymbol{x}) + x_2) = 0 \quad \text{for} \quad \boldsymbol{x} \in \Gamma_C. \tag{1.8}$$

The corresponding friction conditions then take the form:

$$|T_1| \leq \mathfrak{F}g, \quad u_1 \neq 0 \Rightarrow T_1 = -\mathfrak{F}g \operatorname{sgn}(u_1) \quad \text{on} \quad \Gamma_C. \tag{1.9}$$

Convention. Note that (1.8)–(1.9) and (1.5)–(1.6) are equivalent, provided that $\boldsymbol{\nu}(\boldsymbol{x}) = (0, -1)^T \; \forall \boldsymbol{x} \in \Gamma_C$ in the chosen coordinate system, i.e. if Γ_C is affine. Therefore, given any vector field $\boldsymbol{v} : \Gamma_C \to \mathbb{R}^2$, we will sometimes use the more illustrative terms of tangential and normal component for the coordinate functions v_1, v_2, respectively.

Although until the end of this chapter the parameter $\alpha \in \mathcal{U}_{ad}$ will be fixed (unless stated otherwise), the notation shall highlight the fact, that a given quantity depends on this parameter. To this end, we will write e.g. $\Gamma_C(\alpha)$, $\boldsymbol{L}_\alpha$, $\boldsymbol{u}(\alpha)$, etc.

1.1.1 Primal variational formulation

Let us proceed with the weak formulation of (1.1)–(1.4) and (1.8)–(1.9). In order to do so, let us introduce the following function spaces:

$$
\begin{aligned}
V(\alpha) &:= \{v \in H^1(\Omega(\alpha)) \mid v = 0 \text{ on } \Gamma_D(\alpha)\},\\
\boldsymbol{V}(\alpha) &:= V(\alpha) \times V(\alpha),\\
\boldsymbol{K}(\alpha) &:= \{\boldsymbol{v} \in \boldsymbol{V}(\alpha) \mid -v_2 \le d_\alpha \text{ on } \Gamma_C(\alpha)\},
\end{aligned}
$$

where the equality and inequality conditions on parts of the boundary $\partial\Omega(\alpha)$ are meant in the sense of traces, and the distance function d_α is given by (cf. (1.8)):

$$
d_\alpha(\boldsymbol{x}) := \alpha(x_1) \quad \forall \boldsymbol{x} \in \Gamma_C(\alpha).
$$

As usual, the weak formulation of the problem (1.1)–(1.4), (1.8)–(1.9) can be easily derived by multiplying (1.1) by $(\boldsymbol{v} - \boldsymbol{u})$ for some $\boldsymbol{v} \in \boldsymbol{K}(\alpha)$, applying the Green theorem and using the fact that $\boldsymbol{T} \cdot (\boldsymbol{v} - \boldsymbol{u}) = T_1(v_1 - u_1) + T_2(v_2 - u_2)$. In the end, one arrives at the following definition:

Definition 1. By a *weak solution* to the Signorini problem with given friction we mean any function $\boldsymbol{u} := \boldsymbol{u}(\alpha) \in \boldsymbol{K}(\alpha)$ satisfying the following variational inequality:

$$
\left.
\begin{aligned}
&\text{Find } \boldsymbol{u} \in \boldsymbol{K}(\alpha) \text{ such that:}\\
&a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_\alpha(\boldsymbol{v}) - j_\alpha(\boldsymbol{u}) \ge L_\alpha(\boldsymbol{v} - \boldsymbol{u}) \quad \forall \boldsymbol{v} \in \boldsymbol{K}(\alpha),
\end{aligned}
\right\} \tag{$\mathcal{A}(\alpha)$}
$$

where the bilinear form a_α, linear form L_α and convex, proper functional j_α, respectively, are given by:

$$
a_\alpha(\boldsymbol{u}, \boldsymbol{v}) := \int_{\Omega(\alpha)} \mathcal{C}\varepsilon(\boldsymbol{u}) : \varepsilon(\boldsymbol{v})\, d\boldsymbol{x} \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)), \tag{1.10}
$$

$$
L_\alpha(\boldsymbol{v}) := \int_{\Omega(\alpha)} \boldsymbol{F} \cdot \boldsymbol{v}\, d\boldsymbol{x} + \int_{\Gamma_N(\alpha)} \boldsymbol{P} \cdot \boldsymbol{v}\, ds \quad \forall \boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)), \tag{1.11}
$$

$$
j_\alpha(\boldsymbol{v}) := \int_{\Gamma_C(\alpha)} \mathfrak{F} g |v_1|\, ds \quad \forall \boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)). \tag{1.12}
$$

Concerning the regularity of the data, we will assume the following:

(D1) $\boldsymbol{F} \in \boldsymbol{L}^2(\widehat{\Omega})$,

(D2) $\boldsymbol{P} \in \boldsymbol{H}^1(\widehat{\Omega})$,

(D3) $\mathcal{C} = \left(c_{ijkl}\right)_{i,j,k,l=1}^2$, where $c_{ijkl} \in L^\infty(\widehat{\Omega})$ and satisfy the usual symmetry and ellipticity conditions:

$$
\begin{aligned}
&c_{ijkl} = c_{jikl} = c_{klij} \quad \forall i, j, k, l \in \{1, 2\},\\
&\exists C_{ell} > 0 : \ \mathcal{C}\xi : \xi \ge C_{ell}\|\xi\|^2 \quad \forall \xi \in \mathbb{R}^{2\times 2}, \xi^T = \xi.
\end{aligned}
$$

Moreover, to ensure uniform coercivity of a_α on $\boldsymbol{V}(\alpha) \times \boldsymbol{V}(\alpha)$ with respect to $\alpha \in \mathcal{U}_{ad}$, we will assume that

(A4) $\exists \epsilon_D > 0 \ \forall \alpha \in \mathcal{U}_{ad}: \ \text{meas}(\Gamma_D(\alpha)) \geq \epsilon_D$.

The symbols $\mathcal{C}$, $\boldsymbol{F}$, $\boldsymbol{P}$, appearing in (1.10) and (1.11) are then to be understood as restrictions of the mappings declared in (D1)–(D3) onto $\Omega(\alpha)$ and $\Gamma_N(\alpha)$, respectively. Conditions on $\mathfrak{F}$ and g, guaranteeing existence of a weak solution, are specified below.

Theorem 1. *Let $\mathfrak{F} \in L^\infty(\Gamma_C(\alpha))$, $\mathfrak{F} \geq 0$ and $g \in L^2(\Gamma_C(\alpha))$, $g \geq 0$ be given. Then $(\mathcal{A}(\alpha))$ has a unique solution $\boldsymbol{u} \in \boldsymbol{K}(\alpha)$. Moreover, $\boldsymbol{u}$ may be equivalently characterized as the (unique) solution of the variational problem*

$$\left.\begin{array}{ll} \textit{minimize} & \frac{1}{2} a_\alpha(\boldsymbol{v}, \boldsymbol{v}) + j_\alpha(\boldsymbol{v}) - L_\alpha(\boldsymbol{v}) \\ \textit{subj. to} & \boldsymbol{v} \in \boldsymbol{K}(\alpha). \end{array}\right\} \tag{1.13}$$

Proof. See e.g. [22]. □

1.1.2 Mixed variational formulation

Yet another reformulation of $(\mathcal{A}(\alpha))$ (or (1.13)) is the so-called *mixed formulation*, involving Lagrange multipliers for releasing the nonpenetration condition. This way the constrained minimization problem (1.13) can be turned into a saddle-point problem (for more details on the Lagrange multiplier technique in convex optimization the reader is kindly referred to [12]). Before giving the anounced mixed formulation, we will need some more notation to introduce the Lagrange multiplier space:

$$X(\alpha) := \{\varphi \in L^2(\Gamma_C(\alpha)) \mid \exists v \in V(\alpha): \ v = \varphi \text{ on } \Gamma_C(\alpha)\}, \tag{1.14}$$
$$X_+(\alpha) := \{\varphi \in X(\alpha) \mid \varphi \geq 0 \text{ on } \Gamma_C(\alpha)\}, \tag{1.15}$$
$$X'(\alpha) \text{ denotes the topological dual to } X(\alpha), \tag{1.16}$$
$$X'_+(\alpha) := \{\mu \in X'(\alpha) \mid \langle \mu, \varphi \rangle_{X'(\alpha), X(\alpha)} \geq 0 \ \forall \varphi \in X_+(\alpha)\}. \tag{1.17}$$

It can be easily seen that

$$\boldsymbol{v} \in \boldsymbol{K}(\alpha) \ \Leftrightarrow \ \boldsymbol{v} \in \boldsymbol{V}(\alpha) \text{ and } \langle \mu, v_2 + d_\alpha \rangle_{X'(\alpha), X(\alpha)} \geq 0 \ \forall \mu \in X'_+(\alpha).$$

In light of the above characterization of the closed, convex cone $\boldsymbol{K}(\alpha)$, the Lagrangian corresponding to (1.13) is given for each $(\boldsymbol{v}, \mu) \in \boldsymbol{V}(\alpha) \times X'_+(\alpha)$ by

$$\mathcal{L}_\alpha(\boldsymbol{v}, \mu) := \frac{1}{2} a_\alpha(\boldsymbol{v}, \boldsymbol{v}) + j_\alpha(\boldsymbol{v}) - L_\alpha(\boldsymbol{v}) - \langle \mu, v_2 + d_\alpha \rangle_{X'(\alpha), X(\alpha)}.$$

Let us recall that by a *saddle-point* of $\mathcal{L}_\alpha$ we mean a pair $(\boldsymbol{u}, \lambda) \in \boldsymbol{V}(\alpha) \times X'_+(\alpha)$ satisfying:

$$\mathcal{L}_\alpha(\boldsymbol{u}, \mu) \leq \mathcal{L}_\alpha(\boldsymbol{u}, \lambda) \leq \mathcal{L}_\alpha(\boldsymbol{v}, \lambda) \quad \forall (\boldsymbol{v}, \mu) \in \boldsymbol{V}(\alpha) \times X'_+(\alpha).$$

In the context of mixed variational formulations, $\boldsymbol{v}$ is called the primal variable and μ the dual variable, explaining also the title of the previous section.
Concerning the existence of saddle-points of the Lagrangian, we may state following result.

Theorem 2. *Let the assumptions of Theorem 1 be fulfilled. Then $\mathcal{L}_\alpha$ has exactly one saddle-point $(\boldsymbol{u}, \lambda) \in \boldsymbol{K}(\alpha) \times X'_+(\alpha)$, that is also the only solution of:*

$$\left.\begin{array}{l} \text{Find } (\boldsymbol{u}, \lambda) \in \boldsymbol{V}(\alpha) \times X'_+(\alpha) \text{ such that:} \\ a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_\alpha(\boldsymbol{v}) - j_\alpha(\boldsymbol{u}) \\ \qquad\qquad \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}) + \langle \lambda, v_2 - u_2 \rangle_{X'(\alpha), X(\alpha)} \quad \forall \boldsymbol{v} \in \boldsymbol{V}(\alpha), \\ \langle \mu - \lambda, u_2 + d_\alpha \rangle_{X'(\alpha), X(\alpha)} \geq 0 \quad \forall \mu \in X'_+(\alpha). \end{array}\right\} \quad (\bar{\mathcal{A}}(\alpha))$$

Moreover, the first component of the saddle-point satisfies: $\boldsymbol{u} \in \boldsymbol{K}(\alpha)$ and is the unique solution of $(\mathcal{A}(\alpha))$, whereas $\lambda = T_2(\boldsymbol{u})$.

Proof. Existence and uniqueness of the saddle-point follows from [12]; for the second assertion see [22]. □

The saddle-point system $(\bar{\mathcal{A}}(\alpha))$ is called the *mixed formulation* of $(\mathcal{A}(\alpha))$ and it is this formulation we will consider in our state problems, since it allows for the direct computation of the normal contact stress T_2, as well.

1.1.3 Approximation

Now we present a discretization of the contact problems $(\mathcal{A}(\alpha))$ and $(\bar{\mathcal{A}}(\alpha))$ by the finite element method. Throughout this section let the discretization parameter $h := a/(p-1)$, for some $p \in \mathbb{N}, p \geq 2$, be fixed, and denote by $\Delta_h = \{0 = a^1 < a^2 < \cdots < a^p = a\}$ the equidistant partition of $[0, a]$, i.e. $a^i := (i-1)h \ \forall i = 1, \ldots, p$. Let the symbol $\mathbb{P}_1(\Delta_h)$ stand for the set of all piecewise affine functions over Δ_h and let $\alpha_h \in \mathcal{U}^h_{ad} := \mathbb{P}_1(\Delta_h) \cap \mathcal{U}_{ad}$ be given (nonemptiness of $\mathcal{U}^h_{ad}$ is implicitly assumed). On the polygonal domain $\Omega(\alpha_h)$ we introduce a triangulation $\mathcal{T}_h(\alpha_h)$, that meets the following requirements:

(T1) the nodes of $\mathcal{T}_h(\alpha_h)$ lie on the lines $\{a^i\} \times \mathbb{R}$, $i = 1, \ldots, p$ for all $\alpha_h \in \mathcal{U}^h_{ad}$;

(T2) the number of nodes in $\mathcal{T}_h(\alpha_h)$ as well as the neighbours of each triangle from $\mathcal{T}_h(\alpha_h)$ are the *same* for all $\alpha_h \in \mathcal{U}^h_{ad}$;

(T3) the position of nodes of $\mathcal{T}_h(\alpha_h)$ depends *smoothly* on changes of $\alpha_h \in \mathcal{U}^h_{ad}$;

(T4) the triangulations $\mathcal{T}_h(\alpha_h)$ are *compatible* with the decomposition of $\partial\Omega(\alpha_h)$ into $\Gamma_C(\alpha_h)$, $\Gamma_D(\alpha_h)$ and $\Gamma_N(\alpha_h)$ for all $\alpha_h \in \mathcal{U}^h_{ad}$.

The triangulations $\mathcal{T}_h(\alpha_h)$ from the system $\{\mathcal{T}_h(\alpha_h) \mid \alpha_h \in \mathcal{U}^h_{ad}\}$ satisfying (T2)–(T4) are called *topologically equivalent*. On $\mathcal{T}_h(\alpha_h)$ we define the standard, conforming piecewise linear finite element space

$$V_h(\alpha_h) := \mathbb{P}_1(\mathcal{T}_h) \cap V(\alpha_h) = \{v_h \in C(\overline{\Omega}(\alpha_h)) \mid v_h|_K \in P_1(K) \ \forall K \in \mathcal{T}_h(\alpha_h), \\ v_h = 0 \text{ on } \overline{\Gamma}_D(\alpha_h)\}$$

and

$$\boldsymbol{V}_h(\alpha_h) := V_h(\alpha_h) \times V_h(\alpha_h).$$

Let further:

$$\boldsymbol{K}_h(\alpha_h) := \boldsymbol{V}_h(\alpha_h) \cap \boldsymbol{K}(\alpha_h) = \{\boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h) \mid -v_{h2} \leq d_{\alpha_h} \text{ on } \overline{\Gamma}_C(\alpha_h)\},$$

where, for simplicity of presentation, we assume that $\overline{\Gamma}_D(\alpha_h) \cap \overline{\Gamma}_C(\alpha_h) = \emptyset$, i.e. all nodes $\boldsymbol{A}^i = (a^i, \alpha_h(a^i)),\ i = 1, \dots, p$ are contact nodes, for each $\alpha_h \in \mathcal{U}^h_{ad}$. Notice, that since both v_{h2} and d_{α_h} are piecewise linear over the same partition of $\Gamma_C(\alpha_h)$, it holds that $\boldsymbol{K}_h(\alpha_h) \neq \emptyset$. Finally, let $r_h : C(\overline{\Gamma}_C(\alpha_h)) \to \mathbb{P}_1(\mathcal{T}_h(\alpha_h)|_{\Gamma_C(\alpha_h)}) \cap C(\overline{\Gamma}_C(\alpha_h))$ denote the piecewise linear Lagrange interpolation operator on the partition $\mathcal{T}_h(\alpha_h)|_{\Gamma_C(\alpha_h)}$ of $\overline{\Gamma}_C(\alpha_h)$. Now we state the discretized version of our contact problem as follows.

Definition 2. By a *discrete solution* of the Signorini problem with given friction we mean any function $\boldsymbol{u}_h \in \boldsymbol{K}_h(\alpha_h)$ satisfying:

$$\left.\begin{array}{l} \text{Find } \boldsymbol{u}_h \in \boldsymbol{K}_h(\alpha_h) \text{ such that for all } \boldsymbol{v}_h \in \boldsymbol{K}_h(\alpha_h) : \\ a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\boldsymbol{v}_h) - j_{h,\alpha_h}(\boldsymbol{u}_h) \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h), \end{array}\right\} \qquad (\mathcal{A}_h(\alpha_h))$$

where

$$j_{h,\alpha_h}(\boldsymbol{v}_h) := \int_{\Gamma_C(\alpha_h)} \mathfrak{F} g\, r_h |v_{h1}|\, ds \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h). \tag{1.18}$$

Remark 1. The use of the Lagrange interpolation operator in (1.18) might seem unjustified. Nevertheless, it will make more sense in the model with Coulomb's law of friction. At this point just let us note, that the convex functional j_α in (1.12) can be defined with $g \in X'(\alpha)$ as well: $j_\alpha(\boldsymbol{v}) = \langle \mathfrak{F} g, |v_1| \rangle_{X'(\alpha), X(\alpha)}\ \forall \boldsymbol{v} \in \boldsymbol{V}(\alpha)$. Its discretization then involves a functional $g_h \in X'_h(\alpha_h)$, where the definition of the discrete trace space $X_h(\alpha_h)$ is analogous to (1.14) (see also below).

Since the discretization introduced above is conforming, i.e. $\boldsymbol{V}_h(\alpha_h) \subset \boldsymbol{V}(\alpha_h)$ and $\boldsymbol{K}_h(\alpha_h) \subset \boldsymbol{K}(\alpha_h)$, and j_{h,α_h} is convex, lower semicontinuous, the following theorem is obvious (compare with Theorem 1).

Theorem 3. *Let the assumptions of Theorem 1 hold. Then $(\mathcal{A}_h(\alpha_h))$ has exactly one solution $\boldsymbol{u}_h \in \boldsymbol{K}_h(\alpha_h)$. Moreover, $\boldsymbol{u}_h$ is the unique solution of the following convex optimization problem:*

$$\left.\begin{array}{ll} \textit{minimize} & \frac{1}{2} a_{\alpha_h}(\boldsymbol{v}_h, \boldsymbol{v}_h) + j_{h,\alpha_h}(\boldsymbol{v}_h) - L_{\alpha_h}(\boldsymbol{v}_h), \\ \textit{subj. to} & \boldsymbol{v}_h \in \boldsymbol{K}_h(\alpha_h). \end{array}\right\} \tag{1.19}$$

Next we turn to the discretization of the mixed problem $(\bar{\mathcal{A}}(\alpha_h))$. One of the advantages of the mixed variational formulation is that it allows the (almost) independent approximation of the primal and dual variables, i.e. the displacement $\boldsymbol{u}$ and normal contact stress λ in our case. We will present two examples. However, before being able to do so, we need to introduce the discrete counterparts of the spaces (1.14)–(1.17).

Convention. For any function $\varphi \in \Gamma_C(\beta) \to \mathbb{R}$, defined on the contact boundary $\Gamma_C(\beta)$ for some $\beta \in \mathcal{U}_{ad}$, we will denote by $\widehat{\varphi} : (0, a) \to \mathbb{R}$ its transport onto $(0, a)$, i.e. $\widehat{\varphi}(x_1) := \varphi(x_1, \beta(x_1))$ for $x_1 \in (0, a)$.

Keeping this convention in mind, let (compare with (1.14)–(1.15) in the continuous setting):

$$\widehat{X}_h := \{\varphi_h \in L^2(0, a) \mid \exists \alpha_h \in \mathcal{U}^h_{ad}\ \exists v_h \in V_h(\alpha_h) :\ \widehat{v}_h = \varphi_h \text{ in } (0, a)\},$$
$$\widehat{X}_{h+} := \{\varphi_h \in \widehat{X}_h \mid \varphi \geq 0 \text{ in } (0, a)\}.$$

It is easy to see, that $\widehat{X}_h$ is actually independent of α_h and $\widehat{X}_h = \mathbb{P}_1(\Delta_h) \cap C([0,a])$ (recall our assumption that $\overline{\Gamma}_D(\alpha_h) \cap \overline{\Gamma}_C(\alpha_h) = \emptyset$). Notice that in particular $\alpha_h \in \widehat{X}_{h+}$ holds.

The Lagrange multiplier set $X'_+(\alpha)$, defined in (1.17), shall be approximated in the following manner. Let $\widehat{L}_H$ be a finite dimensional space that is in duality with $\widehat{X}_h$ and denote by $\langle \cdot, \cdot \rangle_{Hh} : \widehat{L}_H \times \widehat{X}_h \to \mathbb{R}$ a duality pairing between the two spaces. Then, let $\widehat{\Lambda}_H$ denote the cone of all nonnegative elements of $\widehat{L}_H$, i.e. for each $\mu_H \in \widehat{\Lambda}_H$ it holds that: $\langle \mu_H, \varphi_h \rangle_{Hh} \geq 0 \ \forall \varphi_h \in \widehat{X}_{h+}$. The only requirement concerning $\widehat{L}_H$ we shall need, is the following stability property:

$$\left[\langle \mu_H, \varphi_h \rangle_{Hh} = 0 \quad \forall \varphi_h \in \widehat{X}_h \right] \ \Rightarrow \ \mu_H = 0. \tag{1.20}$$

Now, the discrete Lagrangian on $\boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H$ is given by:

$$\mathcal{L}_{hH,\alpha_h}(\boldsymbol{v}_h, \mu_H) := \frac{1}{2} a_{\alpha_h}(\boldsymbol{v}_h, \boldsymbol{v}_h) + j_{h,\alpha_h}(\boldsymbol{v}_h) - L_{\alpha_h}(\boldsymbol{v}_h) - \langle \mu_H, \widehat{v}_{h2} + \alpha_h \rangle_{Hh}.$$

The next result should be compared to Theorem 2.

Theorem 4. *Let the assumptions of Theorem 1 and the condition* (1.20) *hold true. Then $\mathcal{L}_{hH,\alpha_h}$ has exactly one saddle-point $(\boldsymbol{u}_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H$. It can be determined as the unique solution of the saddle-point system:*

$$\left.\begin{aligned} &\textit{Find } (\boldsymbol{u}_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H \textit{ such that:} \\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\boldsymbol{v}_h) - j_{h,\alpha_h}(\boldsymbol{u}_h) \\ &\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) + \langle \lambda_H, \widehat{v}_{h2} - \widehat{u}_{h2} \rangle_{Hh} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h), \\ &\langle \mu_H - \lambda_H, \widehat{u}_{h2} + \alpha_h \rangle_{Hh} \geq 0 \quad \forall \mu_H \in \widehat{\Lambda}_H. \end{aligned}\right\} \tag{$\bar{\mathcal{A}}_{hH}(\alpha_h)$}$$

Moreover, the first component $\boldsymbol{u}_h$ of the saddle-point is the (unique) solution of the variational inequality:

$$\left.\begin{aligned} &\textit{Find } \boldsymbol{u}_h \in \boldsymbol{K}_{hH}(\alpha_h) \textit{ such that:} \\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\boldsymbol{v}_h) - j_{h,\alpha_h}(\boldsymbol{u}_h) \\ &\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{K}_{hH}(\alpha_h), \end{aligned}\right\} \tag{$\mathcal{A}_{hH}(\alpha_h)$}$$

where:

$$\boldsymbol{K}_{hH}(\alpha_h) := \{ \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h) \mid -\langle \mu_H, \widehat{v}_{h2} \rangle_{Hh} \leq \langle \mu_H, \alpha_h \rangle_{Hh} \ \forall \mu_H \in \widehat{\Lambda}_H \}. \tag{1.21}$$

The second component λ_H of the saddle-point is the Lagrange multiplier releasing the constraint $\boldsymbol{u}_h \in \boldsymbol{K}_{hH}(\alpha_h)$.

Proof. Follows from [12] and (1.20). □

Notice that the "primal" variational inequality $(\mathcal{A}_{hH}(\alpha_h))$ is, in general, different from $(\mathcal{A}_h(\alpha_h))$, as demonstrated by the following two examples.

Example 1. First, we construct a "conforming" discretization of $(\bar{\mathcal{A}}(\alpha))$. To this end, let δ^i denote the Dirac measure concentrated in the ith node of the partition

Δ_h, $i = 1, \dots, p$ and define $\widehat{L}_H := \operatorname{span}\{\delta^1, \dots, \delta^p\}$, endowed with the standard duality pairing. Hence

$$\widehat{\Lambda}_H = \Big\{ \mu_H \in \widehat{L}_H \;\Big|\; \mu_H = \sum_{i=1}^{p} \mu_i \delta^i,\ \mu_i \in \mathbb{R}_+ \ \forall i = 1, \dots, p \Big\}$$

and the stability condition (1.20) holds true. In addition:

$$\begin{aligned} \boldsymbol{K}_{hH}(\alpha_h) &= \{ \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h) \mid -\widehat{v}_{h2}(a^i) \le \alpha_h(a^i) \ \forall i = 1, \dots, p \} \\ &= \boldsymbol{K}_h(\alpha_h) \subset \boldsymbol{K}(\alpha_h). \end{aligned}$$

Therefore, in this case the first component of the solution to $(\bar{\mathcal{A}}_{hH}(\alpha_h))$ is also the unique solution to the primal variational inequality $(\mathcal{A}_h(\alpha_h))$.

Example 2. In this example we consider a more regular approximation for the Lagrange multiplier space, such that $\widehat{\Lambda}_H \subset L^2(0,a)$ holds. Given the equidistant partition Δ_h of $[0,a]$, we construct another partition $\Delta_H := \{0 = a^{1/2} < a^{3/2} < \dots < a^{p+1/2} = a\}$ by setting $a^{i+1/2} := \frac{1}{2}(a^i + a^{i+1})$ $\forall i = 1, \dots, p-1$ as shown in Figure 1.4. Now let $S^j := (a^{j-1/2}, a^{j+1/2})$ and χ^j be the characteristic function of

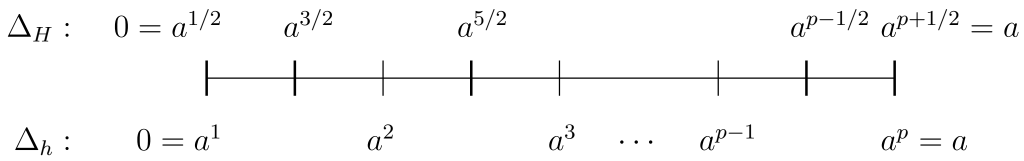


Figure 1.4: Partition Δ_H.

S^j, $j = 1, \dots, p$. Then we define

$$\widehat{L}_H := \mathbb{P}_0(\Delta_H) = \operatorname{span}\{\chi^1, \dots, \chi^p\}$$

to be the space of piecewise constant functions over Δ_H and

$$\widehat{\Lambda}_H = \Big\{ \mu_H \in \widehat{L}_H \;\Big|\; \mu_H = \sum_{i=1}^{p} \mu_i \chi^i,\ \mu_i \in \mathbb{R}_+,\ \forall i = 1, \dots, p \Big\}.$$

Taking the $L^2(0,a)$-scalar product as the duality pairing between $\widehat{L}_H$ and $\widehat{X}_h$, we see that the stability condition (1.20) is satisfied. On the other hand, one gets:

$$\boldsymbol{K}_{hH}(\alpha_h) = \Big\{ \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h) \;\Big|\; -\int_{S^j} \widehat{v}_{h2}\, dx_1 \le \int_{S^j} \alpha_h \, dx_1 \ \forall j = 1, \dots, p \Big\},$$

i.e. the first component of the saddle-point satisfies the nonpenetration condition only in the sense of integral averages. Since $\boldsymbol{K}_{hH}(\alpha_h) \not\subseteq \boldsymbol{K}(\alpha_h)$, we speak of an external approximation of $\boldsymbol{K}(\alpha_h)$.

1.1.4 Algebraic formulation

When deriving the algebraic form of the discretized mixed-type contact problem $(\bar{\mathcal{A}}_{hH}(\alpha_h))$, i.e. in terms of algebraic equations and inequalities, we proceed as follows. First, let us denote $n := \dim \boldsymbol{V}_h(\alpha_h)$ [1] and by construction we have $p = \dim \widehat{X}_h$. Denoting the Lagrange basis functions of these piecewise linear finite element spaces by $\{\varphi_h^1, \dots, \varphi_h^n\}$ and $\{\psi_h^1, \dots, \psi_h^p\}$, respectively, we immediately see that $\boldsymbol{V}_h(\alpha_h)$ is homeomorphic to $\mathbb{R}^n$, and the discrete admissible set $\mathcal{U}_{ad}^h$ can be identified with the convex, compact set

$$U_{ad} := \Big\{ \boldsymbol{\alpha} \in \mathbb{R}_+^p \ \Big| \ 0 \le \alpha_i \le C_0 \ \forall i = 1, \dots, p, \\ |\alpha_i - \alpha_{i+1}| \le C_1 h \ \forall i = 1, \dots, p-1, \\ \frac{2}{h} C_{21} \le \sum_{i=1}^{p-1} (\alpha_i + \alpha_{i+1}) \le \frac{2}{h} C_{22} \Big\}. \tag{1.22}$$

In the sequel, unless stated otherwise, $\boldsymbol{\alpha} \in U_{ad}$ will be fixed and we will consider the mixed problem $(\bar{\mathcal{A}}_{hH}(\alpha_h))$, where $\alpha_h \in \mathcal{U}_{ad}^h$ has the coordinates $\boldsymbol{\alpha}$ with respect to the basis $\{\psi_h^1, \dots, \psi_h^p\}$. For $\boldsymbol{u}_h, \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h)$ denote their coordinates with respect to $\{\varphi_h^1, \dots, \varphi_h^n\}$ by $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$, respectively. The coordinates of $\widehat{v}_{h1}, \widehat{v}_{h2} \in \widehat{X}_h$ with respect to $\{\psi_h^1, \dots, \psi_h^p\}$ shall be denoted in order by $\boldsymbol{v}_\tau \in \mathbb{R}^p$ and $\boldsymbol{v}_\nu \in \mathbb{R}^p$. It is easy to see that $(\boldsymbol{v}_\tau)_i = v_{h1}(a^i, \alpha_h(a^i))$ and $(\boldsymbol{v}_\nu)_i = v_{h2}(a^i, \alpha_h(a^i))$ for every $i = 1, \dots, p$, which means that $\boldsymbol{v}_\tau$ and $\boldsymbol{v}_\nu$ are actually subvectors of $\boldsymbol{v}$. This justifies their notation, and we shall call $\boldsymbol{v}_\tau$ and $\boldsymbol{v}_\nu$ the *tangential* and *normal component* of $\boldsymbol{v}$ along the contact zone.

Further, we will denote by $\mathbb{A}(\boldsymbol{\alpha}) \in \mathbb{R}^{n \times n}$ and $\boldsymbol{L}(\boldsymbol{\alpha}) \in \mathbb{R}^n$ the *stiffness matrix* and *load vector*, respectively, given by

$$\mathbb{A}(\boldsymbol{\alpha}) = \big(a_{\alpha_h}(\varphi_h^i, \varphi_h^j)\big)_{i,j=1,\dots,n} \quad \text{and} \quad \boldsymbol{L}(\boldsymbol{\alpha}) = \big(L_{\alpha_h}(\varphi_h^1), \dots, L_{\alpha_h}(\varphi_h^n)\big)^T.$$

As the triangulations $\mathcal{T}_h(\alpha_h)$ satisfy (T2) and (T3), the mappings $\mathbb{A} : U_{ad} \to \mathbb{R}^{n \times n}$ and $\boldsymbol{L} : U_{ad} \to \mathbb{R}^n$ are *smooth*, i.e. C^1. Moreover, the matrices $\mathbb{A}(\boldsymbol{\alpha})$ are symmetric and uniformly positive definite:

$$\exists \gamma > 0 : \quad \langle \mathbb{A}(\boldsymbol{\alpha}) \boldsymbol{v}, \boldsymbol{v} \rangle_n \ge \gamma \|\boldsymbol{v}\|_n^2 \quad \forall \boldsymbol{v} \in \mathbb{R}^n \ \forall \boldsymbol{\alpha} \in U_{ad}, \tag{1.23}$$

as follows from Korn's inequality and the topological equivalence (T2)–(T4) of the triangulations, ensuring that the mapping $\boldsymbol{\alpha} \mapsto \mathbb{A}(\boldsymbol{\alpha})$ is smooth.

For the evaluation of the nonlinear frictional term j_{h,α_h}, defined by (1.18), we will use numerical integration, namely, the compound rectangle rule over the "refined" partition $\Delta_h \cup \Delta_H$ of $[0, a]$ with $a^i \in \Delta_h$, $i = 1, \dots, p$ as the integration nodes. This means that for every $i = 1, \dots, p$:

$$\int_{a^{i-1/2}}^{a^i} \widehat{\mathfrak{F}} \widehat{g} \pi_h |\widehat{v}_{h1}| \sqrt{1 + (\alpha_h')^2} \, dx_1 \approx \frac{1}{2} \sqrt{h^2 + (\alpha_{i-1} - \alpha_i)^2} \mathfrak{F}_i g_i |(\boldsymbol{v}_\tau)_i|, \tag{1.24}$$

where $\mathfrak{F}_i := \mathfrak{F}(a^i, \alpha_h(a^i))$, $g_i := g(a^i, \alpha_h(a^i))$ and $\pi_h : C([0, a]) \to \mathbb{P}_1(\Delta_h) \cap C([0, a])$ stands for the piecewise linear Lagrange interpolation operator defined

[1] Note, that due to condition (T2), satisfied by the triagulations $\mathcal{T}_h(\alpha_h)$, n is independent of $\alpha_h \in \mathcal{U}_{ad}^h$.

on Δ_h. Obviously: $\widehat{r_h\varphi} = \pi_h\widehat{\varphi}$ for every $\varphi \in C(\overline{\Gamma}_C(\alpha_h))$ and $\alpha_h \in \mathcal{U}^h_{ad}$. Similarly to (1.24) we write the quadrature rule on $[a^i, a^{i+1/2}]$ and sum both expressions for $i = 1, \dots, p$ to obtain:

$$j_{h,\alpha_h}(\boldsymbol{v}_h) \approx \sum_{i=1}^{p} \omega_i(\boldsymbol{\alpha})\mathfrak{F}_i g_i |(\boldsymbol{v}_\tau)_i|, \tag{1.25}$$

where

$$\omega_i(\boldsymbol{\alpha}) = \begin{cases} \frac{1}{2}\sqrt{h^2 + (\alpha_1 - \alpha_2)^2} & \text{if } i = 1, \\ \frac{1}{2}\big(\sqrt{h^2 + (\alpha_{i-1} - \alpha_i)^2} + \sqrt{h^2 + (\alpha_i - \alpha_{i+1})^2}\big) & \text{if } 2 \le i \le p-1, \\ \frac{1}{2}\sqrt{h^2 + (\alpha_{p-1} - \alpha_p)^2} & \text{if } i = p. \end{cases} \tag{1.26}$$

For the discrete Lagrange multiplier space $\widehat{L}_H$ we choose the piecewise constant functions over Δ_H as described in Example 2 of the previous section and apply the same quadrature rule as above to evaluate the terms $\langle \mu_H, \widehat{v}_{h2} \rangle_{Hh}$, $\mu_H \in \widehat{\Lambda}_H$. Denoting the coordinates of μ_H with respect to the basis $\{\chi_{S^1}, \dots, \chi_{S^p}\}$ by $\boldsymbol{\mu} \in \mathbb{R}^p_+$, we have:

$$\langle \mu_H, \widehat{v}_{h2} \rangle_{Hh} \approx \sum_{i=1}^{p} h_i \mu_i (\boldsymbol{v}_\nu)_i,$$

where $h_i := h/2$ if $i = 1$ or $i = p$, and $h_i := h$ otherwise. However, instead of the quantities $\mu_i \ge 0$, we will be computing $(h_i\mu_i) \ge 0$, as follows from the definition of the algebraic problem below.

Definition 3. By the *algebraic* Signorini problem with given friction we mean the following variational inequality:

$$\left.\begin{aligned} &\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+ \text{ such that:} \\ &\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u} \rangle_n + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau| \rangle_p \\ &\qquad\qquad \ge \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u} \rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu \rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n, \\ &\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha} \rangle_p \ge 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p_+, \end{aligned}\right\} \tag{$\bar{A}(\boldsymbol{\alpha})$}$$

where the operator $\bullet : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^p$, $\boldsymbol{u} \bullet \boldsymbol{v} := (u_1v_1, \dots, u_pv_p)^T$, denotes the elementwise product of vectors.

Remark 2. Suppose that $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ is a solution to $(\bar{A}(\boldsymbol{\alpha}))$. Then:

$$\langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu \rangle_p = \sum_{i=1}^{p} \lambda_i (\boldsymbol{v}_\nu)_i = \sum_{i=1}^{p} \frac{\lambda_i}{\omega_i(\boldsymbol{\alpha})} (\boldsymbol{v}_\nu)_i \omega_i(\boldsymbol{\alpha}) \approx \int_{\Gamma_C(\alpha_h)} \lambda_{alg} v_{h2} \, ds,$$

where $\lambda_{alg} \in L^2(\Gamma_C(\alpha_h))$ is such that $\widehat{\lambda}_{alg} = \sum_{i=1}^{p} \frac{\lambda_i}{\omega_i(\boldsymbol{\alpha})} \chi_{S^i} \in \widehat{\Lambda}_H$, i.e. λ_{alg} approximates the Lagrange multiplier $\lambda \in X'_+(\alpha_h)$ from the continuous problem $(\bar{\mathcal{A}}(\alpha_h))$.

When interpreting $\boldsymbol{u}$, the first component of the solution to $(\bar{A}(\boldsymbol{\alpha}))$, we find that $\boldsymbol{u} \in \mathbb{K}(\boldsymbol{\alpha}) = \{\boldsymbol{v} \in \mathbb{R}^n \mid -\boldsymbol{v}_\nu \le \boldsymbol{\alpha}\}$ as follows from $(\bar{A}(\boldsymbol{\alpha}))_2$, which is a consequence of the used integration formula. Moreover, $\boldsymbol{u}$ solves the variational inequality:

$$\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u} \rangle_n + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau| \rangle_p \ge \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u} \rangle_n \quad \forall \boldsymbol{v} \in \mathbb{K}(\boldsymbol{\alpha}).$$

In particular, we see that, as an effect of the numerical integration, one retains an inner approximation $\boldsymbol{K}(\alpha_h)$, i.e. the corresponding discrete solution $\boldsymbol{u}_h$ satisfies the nonpenetration condition along $\Gamma_C(\alpha_h)$. On the other hand, λ_{alg} (see previous Remark 2) and $\boldsymbol{u}_h$ satisfy the complementarity system only approximately.

Remark 3. The algebraic system $(\bar{A}(\boldsymbol{\alpha}))$ is the same as if we had used the approach of Example 1 from the previous section, only the interpretation of the multiplier vector $\boldsymbol{\lambda} \in \mathbb{R}^p_+$ is different–see the explanation above.

For the sake of completeness, let us formulate the following theorem concerning the solvability of $(\bar{A}(\boldsymbol{\alpha}))$.

Theorem 5. *Let $\mathfrak{F}, \boldsymbol{g} \in \mathbb{R}^P_+$ be given. Then $(\bar{A}(\boldsymbol{\alpha}))$ has a unique solution $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ for each $\boldsymbol{\alpha} \in U_{ad}$.*

In the subsequent sections and chapters we shall need the following properties of the solutions to $(\bar{A}(\boldsymbol{\alpha}))$, formulated in a lemma below. Before proceeding to this lemma, however, we give another auxiliary result.

Lemma 1. *There exists a constant $\beta > 0$ such that:*

$$\sup_{\mathbf{0} \neq \boldsymbol{v} \in \mathbb{R}^n} \frac{\langle \boldsymbol{\mu}, \boldsymbol{v}_\nu \rangle_p}{\|\boldsymbol{v}\|_n} \geq \beta \|\boldsymbol{\mu}\|_p \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p. \tag{1.27}$$

Proof. Denote by $\mathbb{N} \in \mathbb{R}^{p \times n}$ the matrix that represents the linear mapping $\boldsymbol{v} \mapsto \boldsymbol{v}_\nu$, i.e. $\mathbb{N}\boldsymbol{v} = \boldsymbol{v}_\nu \ \ \forall \boldsymbol{v} \in \mathbb{R}^n$. Then one has:

$$\sup_{\mathbf{0} \neq \boldsymbol{v} \in \mathbb{R}^n} \frac{\langle \boldsymbol{\mu}, \boldsymbol{v}_\nu \rangle_p}{\|\boldsymbol{v}\|_n} = \sup_{\mathbf{0} \neq \boldsymbol{v} \in \mathbb{R}^n} = \frac{\langle \mathbb{N}^T \boldsymbol{\mu}, \boldsymbol{v} \rangle_n}{\|\boldsymbol{v}\|_n} = \|\mathbb{N}^T \boldsymbol{\mu}\|_n.$$

The function $\boldsymbol{\mu} \mapsto \|\mathbb{N}^T \boldsymbol{\mu}\|_n$ is nonnegative and continuous in $\mathbb{R}^p$, therefore it attains its minimum on the unit sphere. Denoting this minimum value by β, it can be immediately seen that $\beta > 0$ iff $\mathrm{Ker}\,(\mathbb{N}^T) = \{\mathbf{0}\}$, i.e. if $\mathbb{N}$ has full row rank. In our case $\mathbb{N}$ has in each row exactly one element equal to 1, all other elements are 0, and the ones appear at different indices. Thus the proof is complete. ☐

On the basis of the previous result, it is not difficult to derive the following upper bounds on the solution of $(\bar{A}(\boldsymbol{\alpha}))$.

Lemma 2. *(i) Let $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ be the solution to $(\bar{A}(\boldsymbol{\alpha}))$. Then:*

$$\|\boldsymbol{u}\|_n \leq \frac{\|\boldsymbol{L}(\boldsymbol{\alpha})\|_n}{\gamma}, \quad \|\boldsymbol{\lambda}\|_p \leq \frac{1}{\beta}\left(\frac{\|\mathbb{A}(\boldsymbol{\alpha})\|}{\gamma} + 1\right)\|\boldsymbol{L}(\boldsymbol{\alpha})\|_n.$$

(ii) Let, in addition, $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ be the solution to the problem $(\bar{A}(\boldsymbol{\alpha}))$, but with a different load vector $\bar{\boldsymbol{L}} \in \mathbb{R}^n$. Then:

$$\|\boldsymbol{u} - \bar{\boldsymbol{u}}\|_n \leq \frac{\|\boldsymbol{L}(\boldsymbol{\alpha}) - \bar{\boldsymbol{L}}\|_n}{\gamma}, \quad \|\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}}\|_p \leq \frac{1}{\beta}\left(\frac{\|\mathbb{A}(\boldsymbol{\alpha})\|}{\gamma} + 1\right)\|\boldsymbol{L}(\boldsymbol{\alpha}) - \bar{\boldsymbol{L}}\|_n.$$

Proof. See Proposition 3.2 and Proposition 3.5 in [4]. ☐

Let us conclude this section on the classical Signorini problem with Tresca friction by splitting the system of inequalities ($\bar{A}(\boldsymbol{\alpha})$) into separate relations for the "interior" variables, i.e. the degrees of freedom corresponding to nodes lying in the interior of $\Omega(\alpha_h)$ or on $\Gamma_N(\alpha_h)$, and the contact variables, i.e., the ones defined on $\overline{\Gamma}_C(\alpha_h)$. Such a splitting reflects the structure of contact problems more: it shows that the nonsmooth nature of these problems stems from the contact conditions, expressed in terms of variational inequalities for $\boldsymbol{u}_\tau$, $\boldsymbol{u}_\nu$, $\boldsymbol{\lambda}$, whereas the internal variables are linked to the contact ones "only" by means of a linear mapping involving the inverse of a symmetric, positive definite matrix. Finally, we rewrite the resulting system of equations and inequalities into a generalized equation (GE). This form will be more suitable for the techniques involved in sensitivity analysis to come in later chapters.

In order to derive the aforementioned form, we split the displacement fields into two parts: $\boldsymbol{v} = (\boldsymbol{v}_{int}, \boldsymbol{v}_{cont}) \in \mathbb{R}^{n-2p} \times \mathbb{R}^{2p}$, where $\boldsymbol{v}_{cont} = (\boldsymbol{v}_\tau, \boldsymbol{v}_\nu) \in \mathbb{R}^p \times \mathbb{R}^p$ comprises the components of $\boldsymbol{v}$ associated with the tangential and normal displacement along $\Gamma_C(\alpha_h)$. We split the load vector similarly into "interior" and "contact" part: $\boldsymbol{L}(\boldsymbol{\alpha}) = (\boldsymbol{L}_{int}(\boldsymbol{\alpha}), \boldsymbol{L}_{cont}(\boldsymbol{\alpha}))$; the stiffness matrix is handled accordingly:

$$\mathbb{A}(\boldsymbol{\alpha}) = \begin{bmatrix} \mathbb{A}_{ii}(\boldsymbol{\alpha}) & \mathbb{A}_{ic}(\boldsymbol{\alpha}) \\ \mathbb{A}_{ci}(\boldsymbol{\alpha}) & \mathbb{A}_{cc}(\boldsymbol{\alpha}) \end{bmatrix}.$$

Resulting from the properties of $\mathbb{A}(\boldsymbol{\alpha})$, the matrices $\mathbb{A}_{ii}(\boldsymbol{\alpha})$ and $\mathbb{A}_{cc}(\boldsymbol{\alpha})$ are symmetric and uniformly positive definite, whereas $\mathbb{A}_{ic}(\boldsymbol{\alpha}) = \mathbb{A}_{ci}(\boldsymbol{\alpha})^T$. Now, testing the first inequality in ($\bar{A}(\boldsymbol{\alpha})$) by $\boldsymbol{v} = (\boldsymbol{v}_{int}, \boldsymbol{u}_{int})$, $\boldsymbol{v}_{int} \in \mathbb{R}^{n-2p}$ arbitrary, yields:

$$\mathbb{A}_{ii}(\boldsymbol{\alpha})\boldsymbol{u}_{int} = \boldsymbol{L}_{int}(\boldsymbol{\alpha}) - \mathbb{A}_{ic}(\boldsymbol{\alpha})\boldsymbol{u}_{cont}, \tag{1.28}$$

from which:

$$\boldsymbol{u}_{int} = \mathbb{A}_{ii}^{-1}(\boldsymbol{\alpha})\boldsymbol{L}_{int}(\boldsymbol{\alpha}) - \mathbb{A}_{ii}^{-1}(\boldsymbol{\alpha})\mathbb{A}_{ic}(\boldsymbol{\alpha})\boldsymbol{u}_{cont}. \tag{1.29}$$

On the other hand, inserting $\boldsymbol{v} = (\boldsymbol{u}_{int}, \boldsymbol{v}_{cont})$ into ($\bar{A}(\boldsymbol{\alpha})$), such that $\boldsymbol{v}_{cont} \in \mathbb{R}^{2p}$ is arbitrary, gives:

$$\begin{aligned} \langle \mathbb{A}_{ci}(\boldsymbol{\alpha})\boldsymbol{u}_{int} + \mathbb{A}_{cc}(\boldsymbol{\alpha})\boldsymbol{u}_{cont}, \boldsymbol{v}_{cont} - \boldsymbol{u}_{cont}\rangle_{2p} + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ \geq \langle \boldsymbol{L}_{cont}(\boldsymbol{\alpha}), \boldsymbol{v}_{cont} - \boldsymbol{u}_{cont}\rangle_{2p} + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p, \end{aligned}$$

which, combined with (1.29), yields:

$$\begin{aligned} \langle \mathbb{A}_S(\boldsymbol{\alpha})\boldsymbol{u}_{cont}, \boldsymbol{v}_{cont} - \boldsymbol{u}_{cont}\rangle_{2p} + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ \geq \langle \boldsymbol{L}_S(\boldsymbol{\alpha}), \boldsymbol{v}_{cont} - \boldsymbol{u}_{cont}\rangle_{2p} + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p. \end{aligned} \tag{1.30}$$

Here $\mathbb{A}_S(\boldsymbol{\alpha}) := \mathbb{A}_{cc}(\boldsymbol{\alpha}) - \mathbb{A}_{ci}(\boldsymbol{\alpha})\mathbb{A}_{ii}^{-1}(\boldsymbol{\alpha})\mathbb{A}_{ic}(\boldsymbol{\alpha})$ denotes the Schur complement to $\mathbb{A}_{ii}(\boldsymbol{\alpha})$ in $\mathbb{A}(\boldsymbol{\alpha})$ and $\boldsymbol{L}_S(\boldsymbol{\alpha}) := \boldsymbol{L}_{cont}(\boldsymbol{\alpha}) - \mathbb{A}_{ci}(\boldsymbol{\alpha})\mathbb{A}_{ii}^{-1}(\boldsymbol{\alpha})\boldsymbol{L}_{int}(\boldsymbol{\alpha})$.

Further, according to the decomposition $\boldsymbol{v}_{cont} = (\boldsymbol{v}_\tau, \boldsymbol{v}_\nu) \in \mathbb{R}^p \times \mathbb{R}^p$, let us split:

$$\mathbb{A}_S(\boldsymbol{\alpha}) = \begin{bmatrix} \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha}) \\ \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha}) \end{bmatrix},$$

where the submatrices satisfy: $\mathbb{A}_{\tau\tau}(\boldsymbol{\alpha}), \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha}) \in \mathbb{R}^{p\times p}$ are symmetric and uniformly positive definite and $\mathbb{A}_{\nu\tau}(\boldsymbol{\alpha}) = \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})^T$. The vector $\boldsymbol{L}_S(\boldsymbol{\alpha})$ is decomposed analogously into a "tangential" and "normal" part: $\boldsymbol{L}_S(\boldsymbol{\alpha}) = (\boldsymbol{L}_\tau(\boldsymbol{\alpha}), \boldsymbol{L}_\nu(\boldsymbol{\alpha}))$. First, we test (1.30) with $\boldsymbol{v}_{cont} = (\boldsymbol{v}_\tau, \boldsymbol{u}_\nu)$, $\boldsymbol{v}_\tau \in \mathbb{R}^p$ arbitrary and obtain:

$$\langle \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu, \boldsymbol{v}_\tau - \boldsymbol{u}_\tau\rangle_p + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ \geq \langle \boldsymbol{L}_\tau(\boldsymbol{\alpha}), \boldsymbol{v}_\tau - \boldsymbol{u}_\tau\rangle_p \quad \forall \boldsymbol{v}_\tau \in \mathbb{R}^p,$$

or equivalently:

$$\boldsymbol{0} \in \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + \partial j_{\boldsymbol{\alpha}}(\boldsymbol{u}_\tau), \tag{1.31}$$

where $\partial j_{\boldsymbol{\alpha}}$ stands for the convex subdifferential of $j_{\boldsymbol{\alpha}}(\boldsymbol{w}) := \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \boldsymbol{\mathfrak{F}} \bullet \boldsymbol{g}, |\boldsymbol{w}|\rangle_p$, $\boldsymbol{w} \in \mathbb{R}^p$.

In a similar fashion, by inserting $\boldsymbol{v}_{cont} = (\boldsymbol{u}_\tau, \boldsymbol{v}_\nu)$, $\boldsymbol{v}_\nu \in \mathbb{R}^p$ arbitrary into (1.30), we arrive at the equation:

$$\boldsymbol{0} = \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\boldsymbol{\alpha}). \tag{1.32}$$

Finally, employing the notion of the convex normal cone (cf. [12]), we may rewrite the second inequality in $(\bar{A}(\boldsymbol{\alpha}))$, expressing the nonpenetration condition, as:

$$\boldsymbol{0} \in \boldsymbol{u}_\nu + \boldsymbol{\alpha} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}). \tag{1.33}$$

To summarize, we have shown that the pair $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ is a solution to the Signorini problem with Tresca friction $(\bar{A}(\boldsymbol{\alpha}))$ if and only if $\boldsymbol{u} = (\boldsymbol{u}_{int}, \boldsymbol{u}_\tau, \boldsymbol{u}_\nu) \in \mathbb{R}^{n-2p} \times \mathbb{R}^p \times \mathbb{R}^p$, where $\boldsymbol{u}_{int}$ satisfies (1.29) and the contact variables $(\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda})$ solve the following system of GEs:

$$\left.\begin{array}{l} \boldsymbol{0} \in \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + Q_\tau(\boldsymbol{\alpha}, \boldsymbol{u}_\tau), \\ \boldsymbol{0} = \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\boldsymbol{\alpha}), \\ \boldsymbol{0} \in \boldsymbol{u}_\nu + \boldsymbol{\alpha} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}). \end{array}\right\} \tag{1.34}$$

Here the multifunction $Q_\tau : U_{ad} \times \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ is defined as:

$$Q_\tau(\boldsymbol{\alpha}, \boldsymbol{w}) := \partial j_{\boldsymbol{\alpha}}(\boldsymbol{w}) = \partial \sum_{i=1}^{p} \omega_i(\boldsymbol{\alpha})\mathfrak{F}_i g_i |w_i| \\ = \begin{bmatrix} \omega_1(\boldsymbol{\alpha})\mathfrak{F}_1 g_1 \partial|w_1| \\ \vdots \\ \omega_p(\boldsymbol{\alpha})\mathfrak{F}_p g_p \partial|w_p| \end{bmatrix},$$

as follows from the sum rule for the subdifferential of convex functions [12, Proposition 5.6]. Moreover, Theorem 5 ensures that, given any $\boldsymbol{\mathfrak{F}}, \boldsymbol{g} \in \mathbb{R}^p_+$, (1.34) is uniquely solvable for each $\boldsymbol{\alpha} \in U_{ad}$. Once the contact displacements $\boldsymbol{u}_{cont} = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu)$ have been determined, the internal ones $\boldsymbol{u}_{int}$ can be computed by solving the system of linear algebraic equations (1.28).

1.2 Tresca model with a solution-dependent coefficient of fricion

Up to now we have assumed that the coefficient of friction $\mathfrak{F}$ is constant, or depends on the spatial variable only. In some situations, as experiments show, it is more appropriate to model the coefficient of friction as a function of the unknown displacement (or, better, as a function of the slip velocity in dynamic problems) as well. Namely, we will assume that $\mathfrak{F}$ depends on the magnitude of the tangential displacement. The generalized version of the friction condition (1.9) on $\Gamma_C(\alpha)$ now reads as:

$$\left.\begin{array}{l} u_1 = 0 \;\Rightarrow\; |T_1| \le \mathfrak{F}(0)g, \\ u_1 \ne 0 \;\Rightarrow\; T_1 = -\mathfrak{F}(|u_1|)g\,\mathrm{sgn}(u_1) \end{array}\right\} \text{ on } \Gamma_C(\alpha), \tag{1.35}$$

where $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ is *continuous* and *bounded* (for sake of simplicity, we neglect the dependence on the spatial variable); all other symbols have the same meaning as before. As it was done for Definition 1, one may proceed analogously to obtain the weak formulation of (1.1)–(1.4), (1.8) and (1.35):

Definition 4. By a *weak solution* of the Signorini problem with given friction and a solution-dependent coefficient of friction we mean any solution of:

$$\left.\begin{array}{l} \text{Find } \boldsymbol{u} := \boldsymbol{u}(\alpha) \in \boldsymbol{K}(\alpha) \text{ such that:} \\ a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_\alpha(\boldsymbol{u}, \boldsymbol{v}) - j_\alpha(\boldsymbol{u}, \boldsymbol{u}) \ge L_\alpha(\boldsymbol{v} - \boldsymbol{u}) \quad \forall \boldsymbol{v} \in \boldsymbol{K}(\alpha), \end{array}\right\} \tag{$\mathcal{P}(\alpha)$}$$

where

$$j_\alpha(\boldsymbol{w}, \boldsymbol{v}) := \int_{\Gamma_C(\alpha)} \mathfrak{F}(|w_1|)g|v_1|\,ds \quad \forall \boldsymbol{w}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)).$$

Problem $(\mathcal{P}(\alpha))$ is an implicit variational inequality of the second kind. Had we known the function $|u_1| \in X_+(\alpha)$ (cf. (1.15) for the definition of $X_+(\alpha)$) a priori, $(\mathcal{P}(\alpha))$ would turn into a standard variational inequality of the form $(\mathcal{A}(\alpha))$. This trivial observation leads to the following equivalent characterization of the solutions to $(\mathcal{P}(\alpha))$:

Proposition 1. *For any $\varphi \in X_+(\alpha)$ denote the problem $(\mathcal{A}(\alpha))$ with the coefficient of friction $\mathfrak{F} \circ \varphi \in L^\infty(\Gamma_C(\alpha))$ by $(\mathcal{A}(\alpha, \varphi))$. Consider the mapping:*

$$\Phi : X_+(\alpha) \to X_+(\alpha), \quad \varphi \mapsto |u_1(\varphi)_{|\Gamma_C(\alpha)}|,$$

where $\boldsymbol{u}(\varphi)$ is the (unique) solution of $(\mathcal{A}(\alpha, \varphi))$. Then $\boldsymbol{u}$ solves $(\mathcal{P}(\alpha))$ iff $\boldsymbol{u}$ is the solution of $(\mathcal{A}(\alpha, \varphi^*))$, where φ^* is a fixed point of the mapping Φ.

On the basis of Proposition 1 and using appropriate fixed-point theorems, the following existence and uniqueness results are not difficult to prove.

Theorem 6. *Let $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ be continuous and bounded: $\exists C_{max} > 0\ \forall x \in \mathbb{R}_+ : 0 \le \mathfrak{F}(x) \le C_{max}$; $g \in L^2_+(\Gamma_C(\alpha))$. Then $(\mathcal{P}(\alpha))$ has at least one solution.*

Proof. See [19]. □

By strengthening the assumptions on $\mathfrak{F}$ and g, one may ensure unique solvability of $(\mathcal{P}(\alpha))$, as follows from the next theorem.

Theorem 7. *Let, in addition to the assumptions of Theorem 6, $\mathfrak{F}$ be Lipschitz continuous: $\exists C_{lip} > 0\ \forall x, y \in \mathbb{R}_+ : \ |\mathfrak{F}(x) - \mathfrak{F}(y)| \le C_{lip}|x - y|$, and $g \in L^\infty_+(\Gamma_C(\alpha))$. There exists a constant $\bar{C} > 0$ such that if $C_{lip}\|g\|_{L^\infty(\Gamma_C(\alpha))} \in (0, \bar{C})$, then $(\mathcal{P}(\alpha))$ has exactly one solution. Moreover, $\bar{C}$ can be chosen independently of $\alpha \in \mathcal{U}_{ad}$.*

Proof. See [19] and [43, Thm. 1.2], where the exact form of the bound $\bar{C}$ is also given. $\square$

Assuming that the slip bound g is a restriction of a given function $\overline{g} \in C(\overline{\widehat{\Omega}})$ onto $\Gamma_C(\alpha)$, $\|g\|_{L^\infty(\Gamma_C(\alpha))}$ can also be estimated independently of α. In this case the second assertion of the previous theorem yields, that the state problems $(\mathcal{P}(\alpha))$ are uniquely solvable for all $\alpha \in \mathcal{U}_{ad}$, provided the coefficient of friction is Lipschitz continuous with a sufficiently small modulus C_{lip}. In the sequel we will rely on this property when dealing with the shape optimization problem.

Exploiting the fixed-point structure of $(\mathcal{P}(\alpha))$, described in Proposition 1, and using Theorem 2, we may write the mixed form of $(\mathcal{P}(\alpha))$ as follows.

Theorem 8. *Let the assumptions of Theorem 7 hold. Then the system of variational inequalities*

$$\left.\begin{aligned}
&\text{Find } (\boldsymbol{u}, \lambda) \in \boldsymbol{V}(\alpha) \times X'_+(\alpha) \text{ such that:}\\
&a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_\alpha(\boldsymbol{u}, \boldsymbol{v}) - j_\alpha(\boldsymbol{u}, \boldsymbol{u})\\
&\qquad\qquad \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}) + \langle \lambda, v_2 - u_2 \rangle_{X'(\alpha), X(\alpha)} \quad \forall \boldsymbol{v} \in \boldsymbol{V}(\alpha),\\
&\langle \mu - \lambda, u_2 + d_\alpha \rangle_{X'(\alpha), X(\alpha)} \geq 0 \quad \forall \mu \in X'_+(\alpha),
\end{aligned}\right\} \quad (\mathcal{M}(\alpha))$$

has exactly one solution. Moreover, the first component of the solution $\boldsymbol{u}$ *lies in* $\boldsymbol{K}(\alpha)$ *and is the unique solution of* $(\mathcal{P}(\alpha))$*; for the Lagrange multiplier we have:* $\lambda = T_2(\boldsymbol{u})$.

We will call problem $(\mathcal{M}(\alpha))$ the mixed formulation of $(\mathcal{P}(\alpha))$.

1.2.1 Approximation

Instead of discretizing $(\mathcal{P}(\alpha))$ and $(\mathcal{M}(\alpha))$ directly, we define the discrete versions of these problems by means of parametrized Signorini problems with given friction and a coefficient of friction, which *does not* depend on the solution. As in Proposition 1, the value of this parameter will be a fixed-point of a suitable mapping.

Let a discrete design variable $\alpha_h \in \mathcal{U}_{ad}^h$ be given and recall the definition of the finite dimensional spaces $\boldsymbol{V}_h(\alpha_h)$, $\widehat{X}_{h+}$ and $\widehat{\Lambda}_H$. For any $\varphi_h \in \widehat{X}_{h+}$ let us denote the discrete contact problem $(\bar{\mathcal{A}}_{hH}(\alpha_h))$ with the coefficient of friction given by $\mathfrak{F} \circ \varphi_h$ as $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h))$:

$$\left.\begin{aligned}
&\text{Find } (\boldsymbol{u}_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H \text{ such that:}\\
&a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\varphi_h; \boldsymbol{v}_h) - j_{h,\alpha_h}(\varphi_h; \boldsymbol{u}_h)\\
&\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) + \langle \lambda_H, \widehat{v}_{h2} - \widehat{u}_{h2} \rangle_{Hh} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h),\\
&\langle \mu_H - \lambda_H, \widehat{u}_{h2} + d_{\alpha_h} \rangle_{Hh} \geq 0 \quad \forall \mu_H \in \widehat{\Lambda}_H.
\end{aligned}\right\} \quad (\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h))$$

As in (1.18), the functional j_{h,α_h} is defined as:

$$j_{h,\alpha_h}(\varphi_h; \boldsymbol{v}_h) = \int_0^a \mathfrak{F}(\varphi_h) \widehat{g} \pi_h |\widehat{v}_{h1}| \sqrt{1 + (\alpha_h')^2}\, dx_1 \quad \forall \varphi_h \in \widehat{X}_{h+}, \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h),$$

where the first argument of j_{h,α_h} now signifies the composition of $\mathfrak{F}$ with φ_h. Recall, that $\pi_h : C([0, a]) \to \mathbb{P}_1(\Delta_h) \cap C([0, a])$ stands for the piecewise linear Lagrange interpolation operator defined on Δ_h.

Definition 5. Let us define the mapping:

$$\Phi_h : \widehat{X}_{h+} \to \widehat{X}_{h+}, \quad \varphi_h \mapsto \pi_h |\widehat{u}_{h1}(\varphi_h)|,$$

where $\boldsymbol{u}_h(\varphi_h)$ is the solution of $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h))$. Then, by a *discrete solution* to the Signorini problem with given friction and a solution-dependent coefficient of friction we mean the solution $(\boldsymbol{u}_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H$ of $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h^*))$, where φ_h^* is a fixed-point of Φ_h.

Inserting the fixed-point of Φ_h into $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h^*))$, it can be easily seen that $(\boldsymbol{u}_h, \lambda_H)$ is a discrete solution in the sense of Definition 5 whenever $(\boldsymbol{u}_h, \lambda_H)$ solves the following system of variational inequalities:

$$\left.\begin{aligned} &\text{Find } (\boldsymbol{u}_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H \text{ such that:}\\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{v}_h) - j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{u}_h)\\ &\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) + \langle \lambda_H, \widehat{v}_{h2} - \widehat{u}_{h2} \rangle_{Hh} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h),\\ &\langle \mu_H - \lambda_H, \widehat{u}_{h2} + d_{\alpha_h} \rangle_{Hh} \geq 0 \quad \forall \mu_H \in \widehat{\Lambda}_H. \end{aligned}\right\} \qquad (\mathcal{M}_{hH}(\alpha_h))$$

Again, due to Definition 5 via fixed-points and the unique solvability of the auxiliary problems $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h))$ (cf. Theorem 4), quantitative analysis of $(\mathcal{M}_{hH}(\alpha_h))$ can be carried out by means of suitable fixed-point theorems applied to Φ_h. This is the idea behind the proof of the following theorem and can be found in [19].

Theorem 9. *(i) (existence) Let the assumptions of Theorem 6 be satisfied. Then $(\mathcal{M}_{hH}(\alpha_h))$ has at least one solution for each $\alpha_h \in \mathcal{U}_{ad}^h$.*

(ii) (uniqueness) Let the assumptions of Theorem 7 be satisfied, i.e. $\mathfrak{F}$ is bounded and Lipschitz continuous in $\mathbb{R}_+$ with Lipschitz modulus C_{lip}, $g \in L_+^\infty(\Gamma_C(\alpha_h))$. Then there exists a constant $\bar{C}_d > 0$, such that the following implication holds: if $C_{lip}\|g\|_{L^\infty(\Gamma_C(\alpha_h))} < \bar{C}_d$, then $(\mathcal{M}_{hH}(\alpha_h))$ has a unique solution. In addition, the upper bound $\bar{C}_d$ may be chosen independently on h and $\alpha_h \in \mathcal{U}_{ad}^h$.

Remark 4. From [43, Thm. 1.2] and [43, Thm. A.5] it follows that $\bar{C}_d < \bar{C}$, i.e., if C_{lip} is sufficiently small, both the discrete and continuous state problems possess a unique solution.

Finally, the interpretation of the solution $(\boldsymbol{u}_h, \lambda_H)$ to the mixed-type problem $(\mathcal{M}_{hH}(\alpha_h))$ is the following. The first component $\boldsymbol{u}_h \in \boldsymbol{K}_{hH}(\alpha_h) = \{\boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h) \mid -\langle \mu_H, \widehat{v}_{h2} \rangle_{Hh} \leq \langle \mu_H, \alpha_h \rangle_{Hh} \; \forall \mu_H \in \widehat{\Lambda}_H\}$ and solves the following implicit variational inequality:

$$\left.\begin{aligned} &\text{Find } \boldsymbol{u}_h \in \boldsymbol{K}_{hH}(\alpha_h) \text{ such that:}\\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{v}_h) - j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{u}_h)\\ &\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{K}_{hH}(\alpha_h) \end{aligned}\right\}$$

The second component λ_H is the Lagrange multiplier releasing the discretized nonpenetration constraint $\boldsymbol{u}_h \in \boldsymbol{K}_{hH}(\alpha_h)$.

1.2.2 Algebraic form

The algebraic form of $(\mathcal{M}_{hH}(\alpha_h))$, suitable for numerical computations, can now be very easily derived from Definition 5 and the results of Section 1.1.4, where all the necessary work has already been carried out. In the sequel we shall use the notation introduced therein.

Let $\boldsymbol{\alpha} \in U_{ad}$ (cf. (1.22)) be given and fix one $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)^T \in \mathbb{R}^p_+$. By $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ denote the mixed form of the algebraic Signorini problem with given friction $(\bar{A}(\boldsymbol{\alpha}))$, where the coefficient of friction is given by the vector $\mathfrak{F}(\boldsymbol{\varphi}) = \big(\mathfrak{F}(\varphi_1), \dots, \mathfrak{F}(\varphi_p)\big)^T \in \mathbb{R}^p_+$, i.e.

$$\left.\begin{aligned} &\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+ \text{ such that:} \\ &\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ &\qquad\qquad \geq \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n, \\ &\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \alpha\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p_+, \end{aligned}\right\} \qquad (\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$$

Now, as in Definition 5, we define the mapping

$$\Psi_\alpha : \mathbb{R}^p_+ \to \mathbb{R}^p_+, \quad \Psi_\alpha(\boldsymbol{\varphi}) := |\boldsymbol{u}_\tau|, \tag{1.36}$$

where $\boldsymbol{u}_\tau$ is the subvector of the first component of $(\boldsymbol{u}, \boldsymbol{\lambda}) := (\boldsymbol{u}(\boldsymbol{\varphi}), \boldsymbol{\lambda}(\boldsymbol{\varphi}))$, the solution of $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$. Note, that due to Theorem 5, $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ has a unique solution for each $\boldsymbol{\alpha} \in U_{ad}$ and $\boldsymbol{\varphi} \in \mathbb{R}^p_+$, hence Ψ_α is well-defined.

Definition 6. Let $\boldsymbol{\varphi}^*$ be a fixed-point of Ψ_α and $(\boldsymbol{u}^*, \boldsymbol{\lambda}^*)$ the corresponding solution of $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^*))$. Then $(\boldsymbol{u}^*, \boldsymbol{\lambda}^*)$ is called the solution of the *algebraic* Signorini problem with Tresca friction and a solution-dependent coefficient of friction.

As in the continuous and discrete settings, $(\boldsymbol{u}^*, \boldsymbol{\lambda}^*)$ may be equivalently characterized as a solution of the following problem (compare with $(\mathcal{M}(\alpha))$, $(\mathcal{M}_{hH}(\alpha_h))$):

$$\left.\begin{aligned} &\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+ \text{ such that:} \\ &\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \mathfrak{F}(|\boldsymbol{u}_\tau|) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ &\qquad\qquad \geq \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n, \\ &\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha}\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p_+, \end{aligned}\right\} \qquad (M(\boldsymbol{\alpha}))$$

Since $(M(\boldsymbol{\alpha}))$ and $(\mathcal{M}_{hH}(\alpha_h))$ are not equivalent, we will state the unique solvability of $(M(\boldsymbol{\alpha}))$ separately in the next theorem.

Theorem 10. *Let $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ be bounded and Lipschitz continuous in $\mathbb{R}_+$ with modulus $C_{lip} > 0$. There exists a constant $\bar{K} > 0$, independent of $\boldsymbol{\alpha} \in U_{ad}$, such that if $C_{lip}\|\boldsymbol{g}\|_\infty \in (0, \bar{K})$, then $(M(\boldsymbol{\alpha}))$ has exactly one solution for each $\boldsymbol{\alpha} \in U_{ad}$.*

Proof. For the sake of completeness, we include a sketch of the proof. The idea is to show that Ψ_α is contractive, whence the assertion follows immediately by the Banach fixed-point theorem.

Let $\boldsymbol{\varphi}^{(i)} \in \mathbb{R}^p_+$, $i = 1, 2$, be arbitrary and consider the solutions $(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ to $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(i)}))$, $i = 1, 2$. Then

$$\boldsymbol{u}^{(i)} \in \mathbb{K}(\boldsymbol{\alpha}) := \{\boldsymbol{v} \in \mathbb{R}^n \mid -\boldsymbol{v}_\nu \leq \boldsymbol{\alpha}\}, \quad i = 1, 2, \tag{1.37}$$

and solve the respective implicit variational inequalities:

$$\left.\begin{aligned}&\left\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}^{(i)}, \boldsymbol{v}-\boldsymbol{u}^{(i)}\right\rangle_n + \left\langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \mathfrak{F}(\boldsymbol{\varphi}^{(i)}) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau^{(i)}|\right\rangle_p \\ &\quad \geq \left\langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v}-\boldsymbol{u}^{(i)}\right\rangle_n \quad \forall \boldsymbol{v} \in \mathbb{K}(\boldsymbol{\alpha}).\end{aligned}\right\} \qquad (A(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(i)}))$$

Now, test the first inequality $(A(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(1)}))$ with $\boldsymbol{v} := \boldsymbol{u}^{(2)}$, the second inequality $(A(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(2)}))$ by $\boldsymbol{v} := \boldsymbol{u}^{(1)}$ and add both inequalities. After rearranging the terms we get:

$$\begin{aligned}&\left\langle \mathbb{A}(\boldsymbol{\alpha})(\boldsymbol{u}^{(1)}-\boldsymbol{u}^{(2)}), \boldsymbol{u}^{(1)}-\boldsymbol{u}^{(2)}\right\rangle_n \\ &\qquad\qquad \leq \left\langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \left(\mathfrak{F}(\boldsymbol{\varphi}^{(1)}) - \mathfrak{F}(\boldsymbol{\varphi}^{(2)})\right) \bullet \boldsymbol{g}, |\boldsymbol{u}_\tau^{(2)}| - |\boldsymbol{u}_\tau^{(1)}|\right\rangle_p .\end{aligned}$$

Using the Cauchy-Schwarz inequality and the Lipschitz continuity of $\mathfrak{F}$, the right-hand side can be estimated by:

$$\begin{aligned}&\left\langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \left(\mathfrak{F}(\boldsymbol{\varphi}^{(1)}) - \mathfrak{F}(\boldsymbol{\varphi}^{(2)})\right) \bullet \boldsymbol{g}, |\boldsymbol{u}_\tau^{(2)}| - |\boldsymbol{u}_\tau^{(1)}|\right\rangle_p \\ &\qquad \leq \|\boldsymbol{\omega}(\boldsymbol{\alpha})\|_\infty \|\boldsymbol{g}\|_\infty \|\mathfrak{F}(\boldsymbol{\varphi}^{(1)}) - \mathfrak{F}(\boldsymbol{\varphi}^{(2)})\|_p \||\boldsymbol{u}_\tau^{(1)}| - |\boldsymbol{u}_\tau^{(2)}|\|_p \\ &\qquad\qquad \leq \bar{\omega} C_{lip} \|\boldsymbol{g}\|_\infty \|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n,\end{aligned}$$

where $\bar{\omega} := \sup_{\boldsymbol{\alpha} \in U_{ad}} \|\boldsymbol{\omega}(\boldsymbol{\alpha})\|_\infty$. Since $\mathbb{A}(\boldsymbol{\alpha})$ are uniformly positive definite, we get:

$$\gamma \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n^2 \leq \bar{\omega} C_{lip} \|\boldsymbol{g}\|_\infty \|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n. \tag{1.38}$$

Finally:

$$\|\Psi_\alpha(\boldsymbol{\varphi}^{(1)}) - \Psi_\alpha(\boldsymbol{\varphi}^{(2)})\|_p \leq \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n \leq \frac{\bar{\omega}}{\gamma} C_{lip} \|\boldsymbol{g}\|_\infty \|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p$$

and the assertion holds with

$$\bar{K} := \gamma / \bar{\omega}. \tag{1.39}$$

□

1.3 Coulomb model with a solution-dependent coefficient of friction

In this section we will introduce the local Coulomb law of friction, with a coefficient of friction that is already assumed to depend on the solution, as it was the case in the previous section. We will discuss the difference between the Tresca and Coulomb laws of friction, pointing out also the difficulties associated with its analysis. The structure of the present section resembles the previous one's: after giving the definition of our contact problem with Coulomb friction and a solution-dependent coefficient of friction, we shall quickly move on to its discretization and further to the algebraic formulation, that will be of our primary interest. Analogously to the Tresca friction case, conditions guaranteeing unique solvability of the algebraic contact problems with Coulomb friction for each value of the design variable $\boldsymbol{\alpha} \in U_{ad}$ will be given and proved.

Let us return to what we said in the introduction, namely that the Tresca law of friction does not model physical reality well. What are its shortcomings and how can they be overcome? To this end, we take a second look at the friction condition (1.9), in particular its first part, which says that sliding does not occur on the contact boundary, until the tangential stress does not attain a certain activation threshold. Now, the problem is that this threshold function does not distinguish between points that will become (after deformation) in contact and points, that will not. As a consequence, friction forces may act also at points which are not even in contact with the obstacle–evidently, this is physically infeasible. Secondly, one would expect from a reasonable friction condition to take into account also the quality of contact. Namely, it should be in line with our everyday experience that the stronger an object sticks to another, the bigger forces are needed to make them slide. In the Tresca law of friction, the slip bound did not depend on the pressure between the contacting surfaces. Both these deficiencies are remedied by the so-called *Coulomb law* of friction, that is formulated below (taking into account our special geometry):

$$|T_1| \leq \mathfrak{F}(|u_1|)T_2, \quad u_1 \neq 0 \;\Rightarrow\; T_1 = -\operatorname{sgn}(u_1)\mathfrak{F}(|u_1|)T_2 \quad \text{on} \quad \Gamma_C(\alpha). \tag{1.40}$$

Note, that in (1.40) the a-priori given slip bound g is replaced by the unknown normal stress $T_2 := T_2(\boldsymbol{u})$. In the Coulomb law of friction (1.40), if a point $\boldsymbol{x} \in \Gamma_C(\alpha)$ will not be in contact, then $T_2(\boldsymbol{x}) = 0$, as follows from the contact condition (1.8), implying also $T_1(\boldsymbol{x}) = 0$. Moreover, the activation threshold for sliding is in this case an increasing (linear) function of the pressure T_2. Finally, remark that we have already assumed that the coefficient of friction $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ in (1.40) may also depend on the magnitude of the tangential displacement.

Thus, by the *classical solution* to the Signorini problem with Coulomb friction we mean a function $\boldsymbol{u} : \overline{\Omega}(\alpha) \to \mathbb{R}^2$ satisfying the system of equilibrium equations and boundary conditions (1.1)–(1.4), (1.8) and (1.40).

Assuming that the classical solution is sufficiently regular and applying the Green theorem, one may easily derive the variational inequality $\boldsymbol{u}$ satisfies (we will stick to the notation already introduced in Section 1.1):

$$\left.\begin{aligned} &\text{Find } \boldsymbol{u} \in \boldsymbol{K}(\alpha) \text{ such that for every } \boldsymbol{v} \in \boldsymbol{K}(\alpha) : \\ &a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + \int_{\Gamma_C(\alpha)} \mathfrak{F}(|u_1|)T_2(\boldsymbol{u})\big(|v_1| - |u_2|\big)\, ds \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}). \end{aligned}\right\} \tag{1.41}$$

Since $\boldsymbol{u} \in \boldsymbol{K}(\alpha) \subset \boldsymbol{H}^1(\Omega(\alpha))$ only, $T_2(\boldsymbol{u}) \notin L^2(\Gamma_C(\alpha))$ in general, but $T_2(\boldsymbol{u}) \in X'_+(\alpha)$. Therefore, instead of (1.41) one should write:

$$\left.\begin{aligned} &\text{Find } \boldsymbol{u} \in \boldsymbol{K}(\alpha) \text{ such that for every } \boldsymbol{v} \in \boldsymbol{K}(\alpha) \\ &a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + \langle \mathfrak{F}(|u_1|)T_2(\boldsymbol{u}), |v_1| - |u_1| \rangle_{X'(\alpha), X(\alpha)} \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}). \end{aligned}\right\} \tag{$\mathcal{P}^C(\alpha)$}$$

Still, in order to make sense to the duality term, $\mathfrak{F}$ and $\boldsymbol{u}$ should be smooth enough so that $\mathfrak{F}(|u_1|)|v_1| \in X(\alpha)$ $\forall \boldsymbol{v} \in \boldsymbol{V}(\alpha)$ (see [11]). In order to overcome this difficulty, we will assume that $T_2(\boldsymbol{u}) \in L^2(\Gamma_C(\alpha))$ and $\mathfrak{F}$ is *continuous* in $\mathbb{R}_+$. Then the duality may be replaced by the $L^2(\Gamma_C(\alpha))$-scalar product. Such $\boldsymbol{u} \in \boldsymbol{K}(\alpha)$, satisfying $(\mathcal{P}^C(\alpha))$ is called the *weak solution*. In addition, observe that this weak solution may again be characterized as the solution of the auxiliary problem $(\mathcal{A}(\alpha))$ with special coefficient of friction $\mathfrak{F}$ and slip bound g. Therefore, as it was done in Proposition 1, for given $\varphi \in X_+(\alpha)$, $g \in L^2_+(\Gamma_C(\alpha))$ denote by

$(\mathcal{A}(\alpha, \varphi, g))$ the auxiliary problem $(\mathcal{A}(\alpha))$ with the coefficient of friction $\mathfrak{F} \circ \varphi$ and the slip bound g. Further, define the mapping:

$$\Phi^C : X_+(\alpha) \times L^2_+(\Gamma_C(\alpha)) \to X_+(\alpha) \times X'_+(\alpha), \quad (\varphi, g) \mapsto \big(|(u_1)|_{\Gamma_C(\alpha)}|, T_2(\boldsymbol{u})\big),$$

where $\boldsymbol{u} = \boldsymbol{u}(\varphi, g)$ is the unique solution of $(\mathcal{A}(\alpha, \varphi, g))$. Now it easy to see that $\boldsymbol{u} \in \boldsymbol{K}(\alpha)$ solves $(\mathcal{P}^C(\alpha))$ iff it is the solution of $(\mathcal{A}(\alpha, \varphi^*, g^*))$, with $(\varphi^*, g^*) \in X_+(\alpha) \times L^2_+(\Gamma_C(\alpha))$ being a fixed-point of Φ^C.

Observe that the mixed formulation $(\bar{\mathcal{A}}(\alpha))$, which will be denoted $(\bar{\mathcal{A}}(\alpha, \varphi, g))$ in order to stress the dependence on φ and g, is particularly useful here: along with the displacement $\boldsymbol{u}$ we also compute the normal stress $T_2(\boldsymbol{u}) = \lambda$, that may be used for the fixed-point iteration in Φ^C. This motivates us to use the mixed formulation to define the *weak solution.*

Definition 7. By a *weak solution* of the Signorini problem with Coulomb friction and a solution-dependent coefficient of friction we mean the pair $(\boldsymbol{u}, \lambda) \in \boldsymbol{V}(\alpha) \times L^2_+(\Gamma_C(\alpha))$ solving uniquely the mixed problem $(\mathcal{A}(\alpha, \varphi^*, g^*))$, where (φ^*, g^*) is a fixed-point of the mapping:

$$\Phi^C : X_+(\alpha) \times L^2_+(\Gamma_C(\alpha)) \to X_+(\alpha) \times X'_+(\alpha), \quad (\varphi, g) \mapsto \big(|(u_1)|_{\Gamma_C(\alpha)}|, \lambda\big).$$

Equivalently, the weak solution $(\boldsymbol{u}, \lambda)$ satisfies the following mixed-type problem:

$$\left.\begin{aligned} &\text{Find } (\boldsymbol{u}, \lambda) \in \boldsymbol{V}(\alpha) \times L^2_+(\Gamma_C(\alpha)) \text{ such that:} \\ &a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + \langle \mathfrak{F}(|u_1|)\lambda, |v_1| - |u_1| \rangle_{X'(\alpha), X(\alpha)} \\ &\qquad\qquad \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}) + \langle \lambda, v_2 - u_2 \rangle_{X'(\alpha), X(\alpha)} \quad \forall \boldsymbol{v} \in \boldsymbol{V}(\alpha), \\ &\langle \mu - \lambda, u_2 + d_\alpha \rangle_{X'(\alpha), X(\alpha)} \geq 0 \quad \forall \mu \in X'_+(\alpha). \end{aligned}\right\} \qquad (\mathcal{M}^C(\alpha))$$

Remark 5. Unfortunately, the mapping Φ^C is not contractive and therefore the analysis of $(\mathcal{M}^C(\alpha))$ has to be conducted in a different way—we kindly refer to the monograph [11] for some relevant results. As we shall see, such issues are not present in finite dimensions and Definition 7 suits well for the discretization of $(\mathcal{M}^C(\alpha))$.

1.3.1 Approximation

It should come as no surprise that the discretization will once again be based on the fixed-point structure of the problem $(\mathcal{M}^C(\alpha))$, iterating through some discrete mixed problems $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h, g_H))$.

To this end, let $\alpha_h \in \mathcal{U}^h_{ad}$ be fixed and let $\widehat{\Lambda}_H \subset L^2(0, a)$ be as in Example 2, i.e., $\widehat{L}_H$ is the space of all piecewise constant functions over the partition Δ_H and $\widehat{\Lambda}_H$ the cone of its nonnegative elements. Let $\varphi_h \in \widehat{X}_{h+}$, $g_H \in \widehat{\Lambda}_H$ be given and denote by $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h, g_H))$ the discrete mixed formulation of the Signorini problem with given friction g_H and coefficient of friction $\mathfrak{F} \circ \varphi_h$:

$$\left.\begin{aligned} &\text{Find } (u_h, \lambda_H) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H \text{ such that:} \\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\varphi_h, g_H; \boldsymbol{v}_h) - j_{h,\alpha_h}(\varphi_h, g_H; \boldsymbol{u}_h) \\ &\qquad\qquad \geq L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) + \langle \lambda_H, \widehat{v}_{h2} - \widehat{u}_{h2} \rangle_{Hh} \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h(\alpha_h), \\ &\langle \mu_H - \lambda_H, \widehat{u}_{h2} + \alpha_h \rangle_{Hh} \geq 0 \quad \forall \mu_H \in \widehat{\Lambda}_H. \end{aligned}\right\}$$

$$(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h, g_H))$$

Here the functional j_{h,α_h} is now defined as (compare with (1.18)):

$$j_{h,\alpha_h}(\varphi_h, g_H; \boldsymbol{v}_h) := \int_0^a \mathfrak{F}(\varphi_h) g_H \pi_h |\widehat{v}_{h1}| \sqrt{1 + (\alpha_h')^2}\, dx_1.$$

Recall, that π_h stands for the piecewise linear Lagrange interpolation opertor on Δ_h.

Definition 8. By a solution to the *discretized* Signorini problem with Coulomb friction and a solution-dependent coefficient of friction we mean a pair $(\boldsymbol{u}_h^*, \lambda_H^*) \in \boldsymbol{V}_h(\alpha_h) \times \widehat{\Lambda}_H$ solving the problem $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h^*, g_H^*))$, where (φ_h^*, g_H^*) is a fixed-point of the mapping

$$\Phi_h^C : \widehat{X}_{h+} \times \widehat{\Lambda}_H \to \widehat{X}_{h+} \times \widehat{\Lambda}_H, \quad (\varphi_h, g_H) \mapsto \big(\pi_h |\widehat{u}_{h1}|, \lambda_H\big),$$

$(\boldsymbol{u}_h, \lambda_H)$ denoting the solution of $(\bar{\mathcal{A}}_{hH}(\alpha_h, \varphi_h, g_H))$.

Instead of dealing with existence and uniqueness of solutions to the discretized Signorini problem with Coulomb friction and a solution-dependent coefficient of friction (for such result the reader is kindly referred to e.g. [20, Thm. 2.1] and [20, Thm. 2.2]) we immediately proceed with the algebraic formulation. Existence and uniqueness in the fully algebraic setting will be investigated in more detail with corresponding proofs.

1.3.2 Algebraic formulation

Recall that for given $\varphi_h \in \widehat{X}_{h+}$ and $g_H \in \widehat{\Lambda}_H$, with coordinates $\boldsymbol{\varphi} \in \mathbb{R}_+^p$ and $\boldsymbol{g} \in \mathbb{R}_+^p$ with respect to the basis $\{\psi_h^1, \dots, \psi_h^p\}$ and $\{\chi_{S^1}, \dots, \chi_{S^p}\}$, respectively, the algebraic Signorini problem with given friction $\boldsymbol{g}$ and coefficient of friction $\mathfrak{F}(\boldsymbol{\varphi})$ was defined as:

$$\left.\begin{aligned}
&\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^p \text{ such that:}\\
&\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\omega}(\boldsymbol{\alpha}) \bullet \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p\\
&\qquad\qquad \geq \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\\
&\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha}\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}_+^p,
\end{aligned}\right\} \qquad (\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$$

now labeled as $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$ in order to stress the dependence on $(\boldsymbol{\varphi}, \boldsymbol{g})$. In Remark 2 it was shown that the function $\widehat{\lambda}_{alg} = \sum_{i=1}^p \frac{\lambda_i}{\omega_i(\boldsymbol{\alpha})} \chi_{S^i}$ is the approximate Lagrange multiplier, hence it is this function we will use in the definition of Φ_h^C.

Introducing the elementwise division operator $\div : \mathbb{R}^p \times (\mathbb{R} \setminus \{0\})^p \to \mathbb{R}^p$, $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \big(\frac{u_1}{v_1}, \dots, \frac{u_p}{v_p}\big)^T$, we define the mapping

$$\Psi_\alpha^C : \mathbb{R}_+^p \times \mathbb{R}_+^p \to \mathbb{R}_+^p \times \mathbb{R}_+^p, \quad (\boldsymbol{\varphi}, \boldsymbol{g}) \mapsto \big(|\boldsymbol{u}_\tau|, \boldsymbol{\lambda} \div \boldsymbol{\omega}(\boldsymbol{\alpha})\big),$$

where $(\boldsymbol{u}, \boldsymbol{\lambda})$ solves $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Given a fixed-point $(\boldsymbol{\varphi}^*, \boldsymbol{g}^*)$ of Ψ_α^C, the corresponding solution $(\boldsymbol{u}, \boldsymbol{\lambda})$ to $(A(\boldsymbol{\alpha}, \boldsymbol{\varphi}^*, \boldsymbol{g}^*))$ is also a solution to:

$$\left.\begin{aligned}
&\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}_+^p \text{ such that:}\\
&\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \mathfrak{F}(|\boldsymbol{u}_\tau|) \bullet \boldsymbol{\lambda}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p\\
&\qquad\qquad \geq \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\\
&\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha}\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}_+^p.
\end{aligned}\right\} \qquad (M^C(\boldsymbol{\alpha}))$$

We have arrived at the following definition.

Definition 9. By the *algebraic* Signorini problem with Coulomb friction and a solution-dependent coefficient of friction we mean the problem $(M^C(\boldsymbol{\alpha}))$.

Observe that, equivalently, the algebraic solution $(\boldsymbol{u}, \boldsymbol{\lambda})$ can be characterized in yet another way, namely as the (unique) solution to the auxiliary problem:

$$\left.\begin{array}{l}\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+ \text{ such that:} \\ \langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \mathfrak{F}(\boldsymbol{\varphi}^*) \bullet \boldsymbol{g}^*, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p \\ \qquad\qquad \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n, \\ \langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha}\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p_+, \end{array}\right\} \qquad (\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^*, \boldsymbol{g}^*))$$

where $(\boldsymbol{\varphi}^*, \boldsymbol{g}^*)$ is a fixed-point of the mapping

$$\tilde{\Psi}^C_\alpha : (\boldsymbol{\varphi}, \boldsymbol{g}) \mapsto (|\boldsymbol{u}_\tau|, \boldsymbol{\lambda}).$$

The pair $(\boldsymbol{u}, \boldsymbol{\lambda})$ in the definition of $\tilde{\Psi}^C_\alpha$ is the solution of $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$. In other words, one may get the solution of the algebraic contact problem with Coulomb friction by solving the algebraic contact problem with Tresca friction, but *without* $\boldsymbol{\omega}(\boldsymbol{\alpha})$ in the frictional term. Although it may not be apparent at the moment, but the fact that the control parameter $\boldsymbol{\alpha} \in U_{ad}$ is not present in the frictional term of $(M^C(\boldsymbol{\alpha}))$ will make a huge difference when it comes to conducting sensitivity analysis.

Now, let us state and prove the following result on existence and uniqueness of the algebraic solution.

Theorem 11. *(i) Let $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ be continuous and bounded, i.e. $\exists C_{max} > 0\ \forall x \in \mathbb{R}_+ : \quad 0 \leq \mathfrak{F}(x) \leq C_{max}$. Then $(M^C(\boldsymbol{\alpha}))$ has at least one solution for all $\boldsymbol{\alpha} \in U_{ad}$.*

(ii) Let, in addition, $\mathfrak{F}$ be Lipschitz continuous with modulus C_{lip}. If C_{lip} and C_{max} are sufficiently small, all problems $(M^C(\boldsymbol{\alpha}))$ have a unique solution.

Proof. For the sake of this proof, let us define the norm on products of Euclidean spaces $\mathbb{R}^s \times \mathbb{R}^t$ $(s, t \in \mathbb{N})$ by: $\|(\boldsymbol{w}, \boldsymbol{z})\|_{s+t} := \|\boldsymbol{w}\|_s + \|\boldsymbol{z}\|_t$.

(i) Let $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}^p_+ \times \mathbb{R}^p_+$ be given and consider $(\boldsymbol{u}, \boldsymbol{\lambda})$, the unique solution of $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Then, from Lemma 2$(i)$ we have immediately:

$$\|\tilde{\Psi}^C_\alpha(\boldsymbol{\varphi}, \boldsymbol{g})\|_{p+p} \leq \|\boldsymbol{u}\|_n + \|\boldsymbol{\lambda}\|_p \leq \left[\frac{1}{\gamma} + \frac{1}{\beta}\left(\frac{\|\mathbb{A}\|}{\gamma} + 1\right)\right] \|\boldsymbol{L}\| =: R, \qquad (1.42)$$

where $\|\mathbb{A}\| = \sup_{\boldsymbol{\alpha} \in U_{ad}} \|\mathbb{A}(\boldsymbol{\alpha})\|$, $\|\boldsymbol{L}\| = \sup_{\boldsymbol{\alpha} \in U_{ad}} \|\boldsymbol{L}(\boldsymbol{\alpha})\|_n$ and hence R does not depend on $\boldsymbol{\alpha} \in U_{ad}$. Thus, $\tilde{\Psi}^C_\alpha$ maps the ball $\mathbb{B}_R \subset \mathbb{R}^p \times \mathbb{R}^p$ of radius R and center $\mathbf{0}$ into itself. Continuity of $\tilde{\Psi}^C_\alpha$ is very easy to verify: for any convergent sequence $\{(\boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)})\} \subset \mathbb{R}^p_+ \times \mathbb{R}^p_+$, $(\boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)}) \to (\boldsymbol{\varphi}, \boldsymbol{g})$, the sequence of solutions $\{(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)})\}$ to the problems $(\tilde{A}(\alpha, \boldsymbol{\varphi}^{(i)}, \boldsymbol{g}^{(i)}))$, $i = 1, 2, \ldots$, converges to the solution of the limit problem $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Hence, by Brouwer's fixed-point theorem, the assertion follows.

(ii) We proceed analogously as in the case of Tresca friction and show that $\tilde{\Psi}^C_\alpha$ is contractive in $\mathbb{B}_R$. To this end, consider two pairs $(\boldsymbol{\varphi}^i, \boldsymbol{g}^i)$, $\|(\boldsymbol{\varphi}^i, \boldsymbol{g}^i)\|_{p+p} \leq R$, $i = 1, 2$, and follow the steps of the proof of Theorem 10 to get:

$$\gamma\|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_n \leq \langle \mathfrak{F}(\boldsymbol{\varphi}^1) \bullet \boldsymbol{g}^1 - \mathfrak{F}(\boldsymbol{\varphi}^2) \bullet \boldsymbol{g}^2, |\boldsymbol{u}^2_\tau| - |\boldsymbol{u}^1_\tau|\rangle_p,$$

where $(\boldsymbol{u}^1, \boldsymbol{g}^1)$, $(\boldsymbol{u}^2, \boldsymbol{g}^2)$ denote the solutions to $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^1, \boldsymbol{g}^1))$, $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^2, \boldsymbol{g}^2))$, respectively. Using the Cauchy-Schwarz inequality, adding and subtracting the term $\mathfrak{F}(\boldsymbol{\varphi}^1) \bullet \boldsymbol{g}^2$ and making use of the assumptions on $\mathfrak{F}$ we arrive at:

$$\|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_n \leq \frac{1}{\gamma}\Big(C_{max}\|\boldsymbol{g}^1 - \boldsymbol{g}^2\|_p + C_{lip}\|\boldsymbol{g}^2\|_\infty \|\boldsymbol{\varphi}^1 - \boldsymbol{\varphi}^2\|_p\Big) \tag{1.43}$$

$$\leq \frac{1}{\gamma}\max\{C_{max}, RC_{lip}\}\|(\boldsymbol{\varphi}^1 - \boldsymbol{\varphi}^2, \boldsymbol{g}^1 - \boldsymbol{g}^2)\|_{p+p} \tag{1.44}$$

Next, we estimate the difference of the Lagrange multipliers $\boldsymbol{\lambda}^1$, $\boldsymbol{\lambda}^2$. From the first inequality of $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^i, \boldsymbol{g}^i))$ one gets:

$$\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}^i, \boldsymbol{v}\rangle_n = \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v}\rangle_n + \langle \boldsymbol{\lambda}^i, \boldsymbol{v}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\ \boldsymbol{v}_\tau = 0,\ i = 1, 2.$$

Subtracting the two equations from each other yields:

$$\langle \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2, \boldsymbol{v}_\nu\rangle_p = \langle \mathbb{A}(\boldsymbol{\alpha})(\boldsymbol{u}^1 - \boldsymbol{u}^2), \boldsymbol{v}\rangle_n \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\ \boldsymbol{v}_\tau = 0.$$

From Lemma 1, noticing that the supremum can be taken over the whole space $\mathbb{R}^n$, we get:

$$\beta\|\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2\|_p \leq \|\mathbb{A}\|\|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_n.$$

Finally, combining the previous two estimates we obtain:

$$\begin{aligned}\|\tilde{\Psi}^C_\alpha(\boldsymbol{\varphi}^1, \boldsymbol{g}^1) - \tilde{\Psi}^C_\alpha(\boldsymbol{\varphi}^2, \boldsymbol{g}^2)\|_{p+p} &\leq \|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_n + \|\boldsymbol{\lambda}^1 + \boldsymbol{\lambda}^2\|_p \\ &\leq \frac{\beta + \|\mathbb{A}\|}{\beta\gamma}\max\{C_{max}, RC_{lip}\}\|(\boldsymbol{\varphi}^1, \boldsymbol{g}^1) - (\boldsymbol{\varphi}^2, \boldsymbol{g}^2)\|_{p+p}.\end{aligned}$$

Now, the assertion of the theorem holds, provided

$$\max\{C_{max}, RC_{lip}\} < \frac{\beta\gamma}{\beta + \|\mathbb{A}\|}. \tag{1.45}$$

□

Remark 6. We have come to another important difference between the contact problems with Tresca friction $(M(\boldsymbol{\alpha}))$ and Coulomb friction $(M^C(\boldsymbol{\alpha}))$. Namely, that in the latter case the condition guaranteeing unique solvability of the problem for every $\boldsymbol{\alpha} \in U_{ad}$, is *data-dependent*. Indeed, the "constant" R in (1.45) is a continuous, increasing function of $\|\boldsymbol{L}\|$, cf. (1.42). For too large load vectors $\boldsymbol{L}$, (1.45) will become invalid and one may lose uniqueness. However, if (1.45) is satisfied for a given set of data $\mathfrak{F}$ and $\boldsymbol{L}$, then it remains valid for sufficiently small perturbations of the load vector, as well. We shall return to this matter when investigating stability $(\mathcal{M}^C(\alpha))$ with respect to the design variable $\boldsymbol{\alpha} \in U_{ad}$.

Chapter 2

Shape optimization: Tresca case

In the previous chapter we have introduced the Signorini problem and several models of friction, formulated various (not always equivalent) mathematical problems describing this physical phenomena. For a fixed geometry, we were interested in existence and uniqueness of solutions to these problems. In addition, it was shown that the considered problems remain (uniquely) solvable even if the geometry is changed. In the present chapter we focus on the model with Tresca friction where the coefficient of friction depends on the unknown solution and take our considerations to a further level. Namely, we will try to identify an optimal geometry among the set of admissible ones, i.e., find such α^* that the pair $(\alpha^*, \boldsymbol{y}^*)$, where $\boldsymbol{y}^*$ is the solution of the corresponding state problem, minimizes a given cost functional J. After proving existence of at least one optimal shape, we will focus on its identification in practice. As we shall see, a crucial ingredient for an effective numerical solution of the shape optimization problem is the computation of (sub)gradients with respect to the design variable. This is subject of the section covering sensitivity analysis and represents our main results in this chapter.

The structure of the chapter is as follows. We start with the definition of the shape optimization problem for continuous and discretized state problems, recalling results from [43]. Convergence analysis is also treated briefly – for details on these issues we kindly refer to [43] and [17]. For the rest of the chapter we shift our attention entirely to the algebraic state problem $(M(\boldsymbol{\alpha}))$ and the corresponding algebraic shape optimization problem. First, stability analysis is carried out, i.e., it is shown that the state variable is a Lipschitz function of the design. As an easy consequence we obtain existence of a solution to the shape optimization problem. For its numerical solution we employ the implicit programming approach (ImP), which requires computing (sub)gradients of the implicitly defined, nondifferentiable control-to-state mapping. This shall be facilitated by the generalized differential calculus of B. Mordukhovich and the calmness property of the state problem. Most of the results presented in this chapter have been published in the paper [17].

2.1 The continuous and discretized shape optimization problem

As already mentioned above, we will be dealing with the Signorini problem with Tresca friction and a solution-dependent coefficient of friction that was introduced in $(\mathcal{P}(\alpha))$ and reads as:

$$\left.\begin{array}{l}\text{Find } \boldsymbol{u} := \boldsymbol{u}(\alpha) \in \boldsymbol{K}(\alpha) \text{ such that:}\\ a_\alpha(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_\alpha(\boldsymbol{u}, \boldsymbol{v}) - j_\alpha(\boldsymbol{u}, \boldsymbol{u}) \geq L_\alpha(\boldsymbol{v} - \boldsymbol{u}) \quad \forall \boldsymbol{v} \in \boldsymbol{K}(\alpha).\end{array}\right\} \tag{$\mathcal{P}(\alpha)$}$$

Here the bilinear form a_α and linear form L_α are defined by (1.10) and (1.11), respectively, whereas the nonsmooth frictional term has the form

$$j_\alpha(\boldsymbol{w}, \boldsymbol{v}) := \int_{\Gamma_C(\alpha)} \mathfrak{F}(|w_1|) g |v_1| \, ds \quad \forall \boldsymbol{w}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)).$$

The solution of $(\mathcal{P}(\alpha))$ is sought in the closed convex set $\boldsymbol{K}(\alpha) = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega(\alpha)) \mid \boldsymbol{v} = 0 \text{ on } \Gamma_D(\alpha),\ -v_2 \leq d_\alpha \text{ on } \Gamma_C(\alpha)\}$. On the other hand, the admissible set $\mathcal{U}_{ad}$ defined in (1.7) as a subset of Lipschitz functions that are together with their first derivatives equibounded, turns out to be too large. Instead, we will be able to prove existence of an optimal parameter in the following subset of $\mathcal{U}_{ad}$:

$$\widetilde{\mathcal{U}}_{ad} := \Big\{ \alpha \in C^{1,1}([0,a]) \ \Big| \ \begin{array}{l} 0 \leq \alpha \leq C_0 \text{ in } [0,a], \\ |\alpha'| \leq C_1 \text{ in } [0,a], \\ |\alpha''| \leq C_3 \text{ a.e. in } (0,a), \\ C_{21} \leq \displaystyle\int_0^a \alpha(x_1)\, dx_1 \leq C_{22} \Big\}, \end{array} \tag{2.1}$$

i.e. $\widetilde{\mathcal{U}}_{ad}$ contains $C^{1,1}$-functions that have, in addition to (1.7), equibounded second derivatives (they exist a.e. in $(0,a)$ by Rademacher's theorem).

Now, let $J : D \to \mathbb{R}$, where $D := \{(\alpha, \boldsymbol{y}) \mid \alpha \in \widetilde{\mathcal{U}}_{ad},\ \boldsymbol{y} \in \boldsymbol{V}(\alpha)\}$, be a given *cost functional* and denote:

$$\mathcal{G} := \{(\alpha, \boldsymbol{u}) \mid \alpha \in \widetilde{\mathcal{U}}_{ad},\ \boldsymbol{u} \text{ solves } (\mathcal{P}(\alpha))\},$$

i.e., $\mathcal{G}$ is the graph of the *control-to-state mapping* $\widetilde{\mathcal{U}}_{ad} \ni \alpha \mapsto \{\boldsymbol{u} \in \boldsymbol{K}(\alpha) \mid \boldsymbol{u} \text{ solves } (\mathcal{P}(\alpha))\}$ (also called the solution map). Note that the control-to-state map is multivalued, in general.

Definition 10. A domain $\Omega(\alpha^*)$ is called *optimal* iff there exists a $\boldsymbol{u}^* \in \boldsymbol{K}(\alpha^*)$ such that $(\alpha^*, \boldsymbol{u}^*) \in \mathcal{G}$ solves the following problem:

$$\left.\begin{array}{l}\text{Find } (\alpha, \boldsymbol{u}) \in \mathcal{G} \text{ such that:}\\ J(\alpha, \boldsymbol{u}) \leq J(\alpha', \boldsymbol{u}') \quad \forall (\alpha', \boldsymbol{u}') \in \mathcal{G}.\end{array}\right\} \tag{$\mathbf{P}$}$$

$(\mathbf{P})$ is termed the *shape optimization problem.*

In order to establish existence of an optimal domain $\Omega(\alpha^*)$, $\alpha^* \in \widetilde{\mathcal{U}}_{ad}$, we show that $\mathcal{G}$ is (sequentially) compact with respect to a suitable topology τ_D on D. Provided we succeed in finding such τ_D, $(\mathbf{P})$ will have at least one solution for all cost functionals J that are (sequentially) lower semicontinuous with respect to this topology.

Let $\widehat{\boldsymbol{V}} := \boldsymbol{H}^1(\widehat{\Omega})$ and denote by $E_\alpha : \boldsymbol{V}(\alpha) \to \widehat{\boldsymbol{V}}$, $\alpha \in \widetilde{\mathcal{U}}_{ad}$, the continuous, linear extension operator from $\Omega(\alpha)$ into $\widehat{\Omega}$. Further, let $\mathcal{E} : D \to \widetilde{\mathcal{U}}_{ad} \times \widehat{\boldsymbol{V}}$, $(\alpha, \boldsymbol{y}) \mapsto (\alpha, E_\alpha \boldsymbol{y})$. Then $\mathcal{E}$ is injective. Indeed, $(\alpha, E_\alpha \boldsymbol{y}) = (\beta, E_\beta \boldsymbol{z})$ implies $\alpha = \beta$ and, consequently, from $E_\alpha \boldsymbol{y} = E_\alpha \boldsymbol{z}$ one has $\boldsymbol{y} = \boldsymbol{z}$. Let us equip $\widetilde{\mathcal{U}}_{ad}$ with the C^1-topology and $\widehat{\boldsymbol{V}}$ with the weak $\boldsymbol{H}^1$-topology. On $\Sigma := \mathcal{E}(D)$ we consider the relative topology induced by the product topology of $\widetilde{\mathcal{U}}_{ad} \times \widehat{\boldsymbol{V}}$. Now it is easy to see that

$$\tau_D := \{A \subset D \mid \mathcal{E}(A) \text{ is open in } \Sigma\} \tag{2.2}$$

defines a topology on D. Indeed, $\emptyset \in \tau_D$ and since $\mathcal{E} : D \to \Sigma$ is bijective, it preserves set intersections and unions, in particular: $\mathcal{E}(\bigcup A_i) = \bigcup \mathcal{E}(A_i)$ and $\mathcal{E}(A \cap B) = \mathcal{E}(A) \cap \mathcal{E}(B)$ for any subsets $A_i, A, B \subset D$.

Lemma 3. *The set $\mathcal{G}$ is sequentially compact in (D, τ_D), where τ_D is defined by* (2.2)*, i.e., it holds that*

$$\begin{aligned}
&\forall \{(\alpha_n, \boldsymbol{u}_n)\} \subset \mathcal{G}\ \exists \{(\alpha_{n_j}, \boldsymbol{u}_{n_j})\} \subset \{(\alpha_n, \boldsymbol{u}_n)\}\ \exists (\alpha, \boldsymbol{u}) \in \mathcal{G} : \\
&\qquad \alpha_{n_j} \to \alpha \text{ in } C^1([0,a]), \quad E_{\alpha_{n_j}} \boldsymbol{u}_{n_j} \rightharpoonup E_\alpha \boldsymbol{u} \text{ (weakly) in } \boldsymbol{H}^1(\widehat{\Omega}), \quad j \to \infty.
\end{aligned}$$

The proof relies on the fact that the domains $\Omega(\alpha)$, $\alpha \in \widetilde{\mathcal{U}}_{ad}$ have the uniform cone property [6]; thus $\|E_\alpha\|$ may be estimated *independently* of α. At some point in the proof of Lemma 3 (see [17, Lemma 1]) one has to take limit in the frictional term $j_{\alpha_{n_j}}$ as $\alpha_{n_j} \to \alpha$, $j \to \infty$, and here comes into play the additional smoothness requirement in (2.1).

Finally, existence of an optimal domain is merely stated in the next theorem.

Theorem 12. *Let the cost functional J be sequentially lower semicontinuous with respect to the topology τ_D* (2.2)*, i.e.*

$$\left.\begin{aligned}
\alpha_n \to \alpha \quad &\text{in } C^1([0,a]),\ \alpha_n, \alpha \in \widetilde{\mathcal{U}}_{ad}, \\
\boldsymbol{y}_n \rightharpoonup \boldsymbol{y} \quad &\text{in } \boldsymbol{H}^1(\widehat{\Omega}),\ \boldsymbol{y}_n, \boldsymbol{y} \in \boldsymbol{H}^1(\widehat{\Omega})
\end{aligned}\right\} \Longrightarrow \liminf_{n\to\infty} J(\alpha_n, \boldsymbol{y}_n|_{\Omega(\alpha_n)}) \geq J(\alpha, \boldsymbol{y}|_{\Omega(\alpha)}).$$

Then $(\mathbf{P})$ *has at least one solution.*

Next we shortly describe the discretization of the shape optimization problem and present results concerning existence of discrete optimal domains and their convergence to optimal ones in the sense of Definition 10.

Discretization of every shape optimization problem is twofold: on the one hand, the admissible parameter set has to be replaced by a finite dimensional one and, secondly, the state problem has to be discretized. We proceeed analogously as in Section 1.1.3, i.e., define the discretized admissible set by means of piecewise linear functions and approximate $(\mathcal{P}(\alpha))$ using conforming piecewise linear finite elements on a regular triangulation of the corresponding polygonal domain.

Let $h > 0$ be fixed. In contrast to Section 1.1.3 we now have $\mathbb{P}_1(\Delta_h) \cap \widetilde{\mathcal{U}}_{ad} = \emptyset$, because $\widetilde{\mathcal{U}}_{ad}$ contains functions of higher regularity. Therefore, $\widetilde{\mathcal{U}}_{ad}$ has to be approximated in a different manner:

$$\begin{aligned}\widetilde{\mathcal{U}}_{ad}^h := \Big\{\alpha_h \in \mathbb{P}_1(\Delta_h) \cap C([a,b]) \;\Big|\; & 0 \le \alpha_h \le C_0 \;\text{ in }\; [0,a], \\ & |\alpha_h'| \le C_1 \;\text{ a.e. in }\; (0,a), \\ & |\alpha_h(a_{i+1}) - 2\alpha_h(a_i) + \alpha_h(a_{i-1})| \le C_3 h^2 \;\forall i = 2,\dots,p-1, \\ & C_{21} \le \int_0^a \alpha_h(x_1)\,dx_1 \le C_{22}\Big\}.\end{aligned} \tag{2.3}$$

Let $\widetilde{\mathcal{U}}_{ad}^h$ denote the set containing all piecewise linear functions satisfying the constraints in (2.1), but instead of the second derivatives we bound the second finite differences at the nodes of Δ_h. Note that $\widetilde{\mathcal{U}}_{ad}^h \subset \mathcal{U}_{ad}$ (cf. (1.7)), but $\widetilde{\mathcal{U}}_{ad}^h \not\subset \widetilde{\mathcal{U}}_{ad}$, i.e., we have an external approximation of $\widetilde{\mathcal{U}}_{ad}$. For a given $\alpha_h \in \widetilde{\mathcal{U}}_{ad}^h$ we again contruct a triangulation $\mathcal{T}_h(\alpha_h)$ of $\overline{\Omega}(\alpha_h)$ that satisfies (T1)–(T4) and recall the definition of the sets $\boldsymbol{V}_h(\alpha_h), \boldsymbol{K}_h(\alpha_h)$ and the piecewise linear Lagrange interpolation operator $\pi_h : C([0,a]) \to \mathbb{P}_1(\Delta_h) \cap C([0,a])$ from Section 1.1.3 and Section 1.2.1, respectively. Thus the discretized "primal" problem reads as (compare with Section 1.2.1):

$$\left.\begin{aligned}&\text{Find } \boldsymbol{u}_h := \boldsymbol{u}_h(\alpha_h) \in \boldsymbol{K}_h(\alpha_h) \text{ such that:} \\ &a_{\alpha_h}(\boldsymbol{u}_h, \boldsymbol{v}_h - \boldsymbol{u}_h) + j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{v}_h) - j_{h,\alpha_h}(\pi_h|\widehat{u}_{h1}|; \boldsymbol{u}_h) \\ &\qquad \ge L_{\alpha_h}(\boldsymbol{v}_h - \boldsymbol{u}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{K}_h(\alpha_h).\end{aligned}\right\} \tag{$\mathcal{P}_h(\alpha_h)$}$$

The discretized shape optimization problem is defined in a similar way as it was done for $(\mathbf{P})$. To this end, let $D_h := \{(\alpha_h, \boldsymbol{y}_h) \mid \alpha_h \in \widetilde{\mathcal{U}}_{ad}^h,\ \boldsymbol{y}_h \in \boldsymbol{V}_h(\alpha_h)\}$ and denote the graph of the control-to-state mapping associated with $(\mathcal{P}_h(\alpha_h))$ by $\mathcal{G}_h := \{(\alpha_h, \boldsymbol{u}_h) \mid \alpha_h \in \widetilde{\mathcal{U}}_{ad}^h,\ \boldsymbol{u}_h \text{ solves } (\mathcal{P}_h(\alpha_h))\}$. Modifying the approach for the continuous setting appropriately, one finds that $\mathcal{G}_h$ is sequentially compact with respect to the topology τ_{D_h} that is induced by the $(C([0,a]) \times \text{weak-}\boldsymbol{H}^1(\widehat{\Omega}))$-topology on $\widetilde{\mathcal{U}}_{ad}^h \times \boldsymbol{H}^1(\widehat{\Omega})$. For details see Proposition 3.1 in [43]. The following existence theorem is therefore straightforward.

Theorem 13. *Let the cost functional J be sequentially lower semicontinuous in the topology τ_{D_h}, i.e.*

$$\left.\begin{aligned}&\alpha_h^{(n)} \to \alpha_h \;\text{ in } C([0,a]),\; \alpha_h^{(n)}, \alpha_h \in \widetilde{\mathcal{U}}_{ad}^h, \\ &\boldsymbol{y}^{(n)} \rightharpoonup \boldsymbol{y} \;\text{ in } \boldsymbol{H}^1(\widehat{\Omega}),\; \boldsymbol{y}^{(n)}, \boldsymbol{y} \in \boldsymbol{H}^1(\widehat{\Omega})\end{aligned}\right\} \Longrightarrow \liminf_{n\to\infty} J(\alpha_h^{(n)}, \boldsymbol{y}^{(n)}|_{\Omega(\alpha_h^{(n)})}) \ge J(\alpha_h, \boldsymbol{y}|_{\Omega(\alpha_h)}).$$

Then the discretized shape optimization problem:

$$\left.\begin{aligned}&\textit{Find } (\alpha_h, \boldsymbol{u}_h) \in \mathcal{G}_h \textit{ such that:} \\ &J(\alpha_h, \boldsymbol{u}_h) \le J(\alpha_h', \boldsymbol{u}_h') \quad \forall (\alpha_h', \boldsymbol{u}_h') \in \mathcal{G}_h\end{aligned}\right\} \tag{$\mathbf{P}_h$}$$

has at least one solution.

Up to now the discretization parameter $h > 0$ was fixed. In what follows we investigate the relationship between solutions of $(\mathbf{P}_h)$ and $(\mathbf{P})$ as $h \to 0_+$.

First of all, let us note that although α_h are Lipschitz only, by controlling their second finite differences we get: $\alpha_h \in \widetilde{\mathcal{U}}^h_{ad}$ $\forall h$ and $\alpha_h \to \alpha$ in $C([0,a])$, $h \to 0_+$, then $\alpha \in \widetilde{\mathcal{U}}_{ad}$ (cf. [17, Lemma 3]). Moreover, the family $\{\widetilde{\mathcal{U}}^h_{ad} \mid h > 0\}$ is dense in $\widetilde{\mathcal{U}}_{ad}$ with respect to the $C([0,a])$-topology (cf. [17, Lemma 2]). Concerning the cost functional J, this time we shall assume that it is continuous in the following sense:

$$\left.\begin{array}{rll} \alpha_h \to \alpha, & \text{in } C([a,b]), \ \alpha_h \in \widetilde{\mathcal{U}}^h_{ad}, \alpha \in \widetilde{\mathcal{U}}_{ad}, \\ E_{a_h}\boldsymbol{u}_h \rightharpoonup E_\alpha \boldsymbol{u}, & \text{in } \boldsymbol{H}^1(\widehat{\Omega}), \boldsymbol{u}_h, \boldsymbol{u} \text{ solves } (\mathcal{P}_h(\alpha_h)) \text{ and } (\mathcal{P}(\alpha)), \text{ resp.} \end{array}\right\} \Longrightarrow$$
$$\Longrightarrow \lim_{h\to 0_+} J(\alpha_h, \boldsymbol{u}_h) = J(\alpha, \boldsymbol{u}). \tag{2.4}$$

Further, denote:

$$\begin{aligned} \bar{\mathcal{G}} := \{(\alpha, \boldsymbol{u}) \in \mathcal{G} \mid \forall h \to 0_+ \ \exists \{h_j\} \subset \{h\} \ \exists \{(\alpha_{h_j}, \boldsymbol{u}_{h_j})\}, (\alpha_{h_j}, \boldsymbol{u}_{h_j}) \in \mathcal{G}_{h_j} : \\ \alpha_{h_j} \to \alpha \text{ in } C([a,b]) \ \text{ and } \ E_{\alpha_{h_j}} \boldsymbol{u}_{h_j} \rightharpoonup E_\alpha \boldsymbol{u} \text{ in } \boldsymbol{H}^1(\widehat{\Omega}), \ h_j \to 0_+\}. \end{aligned}$$

Then the following convergence result holds.

Theorem 14. *Let J satisfy* (2.4) *and $\{(\alpha^*_h, \boldsymbol{u}^*_h)\}$, $h \to 0_+$, be a sequence of discrete optimal pairs, i.e., $(\alpha^*_h, \boldsymbol{u}^*_h) \in \mathcal{G}_h$ is a solution to $(\mathbf{P}_h)$ for every $h > 0$. Then there exists a subsequence $\{h_j\} \subset \{h\}$ and functions $\alpha^* \in \widetilde{\mathcal{U}}_{ad}$, $\boldsymbol{u}^* \in \boldsymbol{H}^1(\widehat{\Omega})$ such that:*

$$\alpha^*_{h_j} \to \alpha^* \text{ in } C([0,a]), \qquad E_{\alpha_{h_j}} \boldsymbol{u}^*_{h_j} \rightharpoonup \boldsymbol{u}^* \text{ in } \boldsymbol{H}^1(\widehat{\Omega}), \ h_j \to 0_+,$$

and $(\alpha^*, \boldsymbol{u}^*|_{\Omega(\alpha^*)}) \in \mathcal{G}$ satisfies:

$$J(\alpha^*, \boldsymbol{u}^*|_{\Omega(\alpha^*)}) \le J(\bar{\alpha}, \bar{\boldsymbol{u}}) \quad \forall (\bar{\alpha}, \bar{\boldsymbol{u}}) \in \bar{\mathcal{G}}.$$

In addition, if $(\mathcal{P}(\alpha))$ are uniquely solvable for all $\alpha \in \widetilde{\mathcal{U}}_{ad}$, then $\bar{\mathcal{G}} = \mathcal{G}$ and $(\alpha^*, \boldsymbol{u}^*|_{\Omega(\alpha^*)})$ is optimal in the sense of Definition 10.

The set $\bar{\mathcal{G}}$ represents those optimal pairs $(\alpha, \boldsymbol{u}) \in \mathcal{G}$ that can be approximated by a subsequence $\{(\alpha_{h_j}, \boldsymbol{u}_{h_j})\}$ of discrete optimal pairs. Theorem 14 then states that from a sequence of discrete optimal pairs one can always extract a subsequence converging to a generally *sub-optimal* pair $(\alpha^*, \boldsymbol{u}^*|_{\Omega(\alpha^*)}) \in \bar{\mathcal{G}}$, i.e. the optimal one with respect to $\bar{\mathcal{G}}$. Optimality (in the sense of Definitiion 10) is ensured whenever the continuous state problems $(\mathcal{P}(\alpha))$ are uniquely solvable. By Theorem 8 we know that this holds true provided the coefficient of friction $\mathfrak{F}$ is Lipschitz with a sufficiently small modulus.

2.2 The algebraic shape optimization problem

From now on we shall be dealing with the numerical solution of one shape optimization problem $(\mathbf{P}_h)$, therefore let $h > 0$ be fixed in the sequel.

It was already mentioned in Chapter 1 that the presence of the nonlinear frictional term j_{h,α_h} in the state problem makes $(\mathbf{P}_h)$ unsuitable for direct numerical realization. To overcome this, using numerical integration, we transformed $(\mathcal{M}_{hH}(\alpha_h))$ into a system of algebraic inequalities $(M(\boldsymbol{\alpha}))$ in Section 1.2.2. Based on $(M(\boldsymbol{\alpha}))$ we will now formulate the algebraic shape optimization problem as a Mathematical Program with Equilibrium Constraints (MPEC) and employ the Implicit Programming (ImP) technique for its solution.

First, let us notice that the discrete admissible set (2.3) may be identified with the set

$$\begin{aligned}\widetilde{U}_{ad} := \Big\{\boldsymbol{\alpha} \in \mathbb{R}^p_+ \,\Big|\, & 0 \leq \alpha_i \leq C_0 \quad \forall i = 1, \dots, p, \\ & |\alpha_i - \alpha_{i+1}| \leq C_1 h \quad \forall i = 1, \dots, p-1, \\ & |\alpha_{i-1} - 2\alpha_i + \alpha_{i+1}| \leq C_3 h^2 \quad \forall i = 2, \dots, p-1, \\ & \frac{2}{h} C_{21} \leq \sum_{i=1}^{p-1} (\alpha_i + \alpha_{i+1}) \leq \frac{2}{h} C_{22} \Big\}.\end{aligned} \tag{2.5}$$

Actually, in the forthcoming analysis we will only need that $\emptyset \neq \widetilde{U}_{ad} \subset \mathbb{R}^p_+$ is compact and convex.

Next, we simplify our presentation by considering the *reduced* state problem (see (1.34) for the Signorini problem with given friction), i.e., we assume that the cost functional depends only on the contact variables $\boldsymbol{u}_\tau$, $\boldsymbol{u}_\nu$, $\boldsymbol{\lambda}$. If this was not the case, one had to compute sensitivities of $\boldsymbol{u}_{int}$ from (1.29), as well. Nevertheless, using appropriate sum rules, this can be done in a straightforward way and won't be considered here.

For a given $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$, the *reduced* algebraic Signorini problem with Tresca friction and a solution-dependent coefficient, formulated as a system of GEs, reads as follows:

$$\left.\begin{aligned} &\mathbf{0} \in \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + \widetilde{Q}_\tau(\boldsymbol{\alpha}, \boldsymbol{u}_\tau), \\ &\mathbf{0} = \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\boldsymbol{\alpha}), \\ &\mathbf{0} \in \boldsymbol{u}_\nu + \boldsymbol{\alpha} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}), \end{aligned}\right\} \tag{2.6}$$

where the multifunction $\widetilde{Q}_\tau : \widetilde{U}_{ad} \times \mathbb{R}^p_+ \rightrightarrows \mathbb{R}^p_+$ takes the form:

$$\widetilde{Q}_\tau(\boldsymbol{\alpha}, \boldsymbol{w}) = \begin{bmatrix} \omega_1(\boldsymbol{\alpha})\mathfrak{F}(|w_1|)g_1\partial|w_1| \\ \vdots \\ \omega_p(\boldsymbol{\alpha})\mathfrak{F}(|w_p|)g_p\partial|w_p| \end{bmatrix}, \quad \boldsymbol{\alpha} \in \widetilde{U}_{ad}, \; \boldsymbol{w} \in \mathbb{R}^p_+. \tag{2.7}$$

Indeed, it suffices to rewrite the auxiliary problem $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ into the form (1.34) and insert the fixed-point of the mapping Ψ_α, defined in (1.36), which leads directly to (2.6) and (2.7).

In order to write (2.6) in a compact form, we introduce the following notation: the state variable shall be denoted by[1] $\boldsymbol{y} := (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda}) \in \mathbb{R}^{3p}$ and we define the single-valued mapping $\boldsymbol{F} : \widetilde{U}_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}^{3p}$ by

$$\boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) := \mathcal{A}(\boldsymbol{\alpha})\boldsymbol{y} - \boldsymbol{l}(\boldsymbol{\alpha}), \quad (\boldsymbol{\alpha}, \boldsymbol{y}) \in \widetilde{U}_{ad} \times \mathbb{R}^{3p}, \tag{2.8}$$

[1] Actually, $\boldsymbol{y} \in \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p$; nevertheless, for brevity, we shall identify this set with $\mathbb{R}^{3p}$.

where

$$\mathcal{A}(\boldsymbol{\alpha}) := \begin{bmatrix} \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha}) & \mathbf{0} \\ \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha}) & -\mathbb{E} \\ \mathbf{0} & \mathbb{E} & \mathbf{0} \end{bmatrix}, \quad \boldsymbol{l}(\boldsymbol{\alpha}) := \begin{bmatrix} \boldsymbol{L}_{\tau}(\boldsymbol{\alpha}) \\ \boldsymbol{L}_{\nu}(\boldsymbol{\alpha}) \\ -\boldsymbol{\alpha} \end{bmatrix}, \tag{2.9}$$

and, finally, the closed-graph multifunction $\widetilde{Q} : \widetilde{U}_{ad} \times \mathbb{R}^{3p} \rightrightarrows \mathbb{R}^{3p}$ as

$$\widetilde{Q}(\boldsymbol{\alpha}, \boldsymbol{y}) := \begin{bmatrix} \widetilde{Q}_{\tau}(\boldsymbol{\alpha}, \boldsymbol{u}_{\tau}) \\ \mathbf{0} \\ N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}) \end{bmatrix}, \quad \boldsymbol{\alpha} \in \widetilde{U}_{ad}, \ \boldsymbol{y} = (\boldsymbol{u}_{\tau}, \boldsymbol{u}_{\nu}, \boldsymbol{\lambda}) \in \mathbb{R}^{3p}. \tag{2.10}$$

Let us recall that, due to the assumptions (T1)–(T4), $\boldsymbol{F}$ is continuously differentiable. Thus (2.6) may be equivalently rewritten as

$$\left.\begin{array}{l} \text{Find } \boldsymbol{y} \in \mathbb{R}^{3p} \text{ such that:} \\ \mathbf{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + \widetilde{Q}(\boldsymbol{\alpha}, \boldsymbol{y}). \end{array}\right\} \tag{$GE(\boldsymbol{\alpha})$}$$

With $(GE(\boldsymbol{\alpha}))$ we associate the *control-to-state mapping* (solution map) $\widetilde{S} : \widetilde{U}_{ad} \rightrightarrows \mathbb{R}^{3p}$, defined as

$$\widetilde{S}(\boldsymbol{\alpha}) := \{\boldsymbol{y} \in \mathbb{R}^{3p} \mid \mathbf{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + \widetilde{Q}(\boldsymbol{\alpha}, \boldsymbol{y})\}, \tag{2.11}$$

i.e., $\widetilde{S}$ assigns to each design variable $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$ the set of solutions to the (reduced) Signorini problem with Tresca friction and a solution-dependent coefficient of friction $(GE(\boldsymbol{\alpha}))$.

Employing notation from the previous sections, we define the algebraic *shape optimization problem* as

$$\left.\begin{array}{ll} \text{minimize} & J(\boldsymbol{\alpha}, \boldsymbol{y}), \\ \text{subj. to} & \boldsymbol{y} \in \widetilde{S}(\boldsymbol{\alpha}), \\ & \boldsymbol{\alpha} \in \widetilde{U}_{ad}, \end{array}\right\} \tag{$\mathbb{P}$}$$

where $J : \widetilde{U}_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}$ is a given cost functional. $(\mathbb{P})$ is in the form of an MPEC, since it represents an optimization problem where one of the constraints is an equilibrium problem. The main result of this section follows next.

Theorem 15. *Let the assumptions of Theorem 10 be satisfied and $J : \widetilde{U}_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}$ be lower semicontinuous. Then $(\mathbb{P})$ has at least one solution.*

Notice, that $(\mathbb{P})$ may be written as

$$\left.\begin{array}{ll} \text{minimize} & J(\boldsymbol{\alpha}, \boldsymbol{y}), \\ \text{subj. to} & (\boldsymbol{\alpha}, \boldsymbol{y}) \in \widetilde{\mathcal{G}}, \end{array}\right\}$$

where $\widetilde{\mathcal{G}} := \operatorname{Gr} \widetilde{S}$ is the graph of the control-to-state mapping $\widetilde{S}$. Hence, it is sufficient to show that $\widetilde{\mathcal{G}} \subset \mathbb{R}^{4p}$ is *compact*, which in turn immediately yields the assertion of the theorem. In the next section we show that under similar assumptions to that of Theorem 10 (Lipschitz continuity of $\mathfrak{F}$) $\widetilde{S}$ is single-valued and Lipschitzian in the compact domain $\widetilde{U}_{ad}$, implying that its graph is compact in $\mathbb{R}^{4p}$.

2.3 Lipschitzian stability

The main result of this section is to show that the control-to-state mapping $\widetilde{S}$ (cf. (2.11)) is Lipschitz provided the friction coefficient $\mathfrak{F}$ is Lipschitz with a sufficiently small modulus. In other words, we prove that the (unique) solution $(\boldsymbol{u}(\boldsymbol{\alpha}), \boldsymbol{\lambda}(\boldsymbol{\alpha})) \in \mathbb{R}^n \times \mathbb{R}^p_+$ of $(M(\boldsymbol{\alpha}))$ is Lipschitz as a function of the design parameter $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$. In addition, we prove another stability result, namely, that—for fixed $\boldsymbol{\alpha}$—the solution of $(M(\boldsymbol{\alpha}))$ is Lipschitzian with respect to the load vector $\boldsymbol{L} \in \mathbb{R}^n$. This fact will be used later when conducting sensitivity analysis.

First, we provide the following auxiliary result, showing the Lipschitzian stability of the solution to the problem $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ with respect to $\boldsymbol{\varphi}$.

Lemma 4. *Let $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$ be fixed and $\mathfrak{F} : \mathbb{R}_+ \to \mathbb{R}_+$ Lipschitz in $\mathbb{R}_+$ with modulus $C_{lip} > 0$. Then there exists a constant $q > 0$, independent of $\boldsymbol{\alpha}$, such that*

$$\|(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}) - (\boldsymbol{u}^{(2)}, \boldsymbol{\lambda}^{(2)})\|_{n+p} \leq q\|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p \quad \forall \boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)} \in \mathbb{R}^p_+,$$

where $(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ denote the (unique) solution to $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(i)}))$, $i = 1, 2$. Moreover, $q = \kappa C_{lip}$ for some positive constant $\kappa > 0$.

Proof. In the proof of Theorem 10 we have already shown the first part of the assertion (cf. (1.38)), namely

$$\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n \leq \frac{\bar{\omega}}{\gamma}\|\boldsymbol{g}\|_\infty C_{lip}\|\boldsymbol{\varphi}^{(1)} - \boldsymbol{\varphi}^{(2)}\|_p. \tag{2.12}$$

Here we show that a similar estimate holds for the Lagrange multipliers. To this end, test the first inequality of $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(i)}))$ in order by $\boldsymbol{v} := \boldsymbol{0}, 2\boldsymbol{u}^{(i)}$ to see that $\boldsymbol{\lambda}^{(i)}$ $(i = 1, 2)$ satisfies

$$\left\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}^{(i)}, \boldsymbol{v}\right\rangle_n = \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v}\rangle_n + \left\langle \boldsymbol{\lambda}^{(i)}, \boldsymbol{v}_\nu \right\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\ \boldsymbol{v}_\tau = \boldsymbol{0}.$$

Subtracting the two equations for $i = 1, 2$ yields

$$\left\langle \boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}, \boldsymbol{v}_\nu \right\rangle_p = \left\langle \mathbb{A}(\boldsymbol{\alpha})\left(\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\right), \boldsymbol{v}\right\rangle_n \quad \forall \boldsymbol{v} \in \mathbb{R}^n, \boldsymbol{v}_\tau = \boldsymbol{0}.$$

Dividing it by $\|\boldsymbol{v}\|_n$ and taking supremum over the set $\mathcal{S} := \{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{v}_\nu \neq \boldsymbol{0},$ the remaining components of $\boldsymbol{v}$ are $0\}$ we arrive at

$$\|\boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}\|_p = \sup_{\boldsymbol{v} \in \mathcal{S}} \frac{\left\langle \boldsymbol{\lambda}^{(1)} - \boldsymbol{\lambda}^{(2)}, \boldsymbol{v}_\nu \right\rangle_p}{\|\boldsymbol{v}\|_n} \leq \|\mathbb{A}\|\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_p, \tag{2.13}$$

where $\|\mathbb{A}\| := \sup_{\boldsymbol{\alpha} \in \widetilde{U}_{ad}} \|\mathbb{A}(\boldsymbol{\alpha})\|$. Combining (2.13) with (2.12) we find that the assertion of the lemma holds with

$$q := \frac{(\|\mathbb{A}\| + 1)\bar{\omega}}{\gamma}\|\boldsymbol{g}\|_\infty C_{lip}. \tag{2.14}$$

$\square$

2.3.1 Stability with respect to the design variable

Now, we let $\boldsymbol{\varphi}$ be fixed and start with investigating Lipschitzian stability of the solution to $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ with respect to $\boldsymbol{\alpha}$.

Lemma 5. *Let $\boldsymbol{\varphi} \in \mathbb{R}^p_+$ be fixed. Then there exists a constant $c > 0$, which does not depend on $\boldsymbol{\varphi}$ and satisfies*

$$\|(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}) - (\boldsymbol{u}^{(2)}, \boldsymbol{\lambda}^{(2)})\|_{n+p} \leq c\|\boldsymbol{\alpha}^{(1)} - \boldsymbol{\alpha}^{(2)}\|_p \quad \forall \boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)} \in \widetilde{U}_{ad},$$

where $(\boldsymbol{u}^{(i)}, \boldsymbol{\lambda}^{(i)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ stands for the solution of $(\bar{A}(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\varphi}))$, $i = 1, 2$.

Proof. First, we will estimate the difference of the primal variables $\|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n$ using the primal formulation of $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$. Recall from the proof of Theorem 10 (cf. (1.37) and $(A(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(i)})))$ that

$$\boldsymbol{u}^{(i)} \in \mathbb{K}(\boldsymbol{\alpha}^{(i)}) = \{\boldsymbol{v} \in \mathbb{R}^n \mid -\boldsymbol{v}_\nu \leq \boldsymbol{\alpha}^{(i)}\}, \quad i = 1, 2,$$

and they solve the respective variational inequalities:

$$\left.\begin{aligned} &\left\langle \mathbb{A}(\boldsymbol{\alpha}^{(i)})\boldsymbol{u}^{(i)}, \boldsymbol{v} - \boldsymbol{u}^{(i)} \right\rangle_n + \left\langle \boldsymbol{\omega}(\boldsymbol{\alpha}^{(i)}) \bullet \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}^{(i)}_\tau| \right\rangle_p \\ &\quad \geq \left\langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}^{(i)} \right\rangle_n \quad \forall \boldsymbol{v} \in \mathbb{K}(\boldsymbol{\alpha}^{(i)}) \end{aligned}\right\} \qquad (A(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\varphi}))$$

for $i = 1, 2$. Observe that the sets $\mathbb{K}(\boldsymbol{\alpha}^{(i)})$ may be written in the following way:

$$\mathbb{K}(\boldsymbol{\alpha}^{(i)}) = \boldsymbol{a}^{(i)} + \mathbb{K}(\boldsymbol{0}),$$

where the vectors $\boldsymbol{a}^{(i)} \in \mathbb{R}^n$ are such that $\boldsymbol{a}^{(i)}_\nu = -\boldsymbol{\alpha}^{(i)}$ and all its other components are zero. Thus for each $i \in \{1, 2\}$

$$\exists \boldsymbol{w}^{(i)} \in \mathbb{K}(\boldsymbol{0}) : \ \boldsymbol{u}^{(i)} = \boldsymbol{a}^{(i)} + \boldsymbol{w}^{(i)}.$$

Inserting now $\boldsymbol{v} := \boldsymbol{a}^{(i)} + \boldsymbol{w}^{(j)}$ into $(A(\boldsymbol{\alpha}^{(i)}, \boldsymbol{\varphi}))$, $i, j \in \{1, 2\}$, $i \neq j$, and adding the two inequalities yields:

$$\begin{aligned} &\left\langle \mathbb{A}(\boldsymbol{\alpha}^{(1)})(\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}), \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)} \right\rangle_n \\ &\quad \leq \left\langle \mathbb{A}(\boldsymbol{\alpha}^{(1)})(\boldsymbol{a}^{(1)} - \boldsymbol{a}^{(2)}), \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)} \right\rangle_n + \left\langle \left(\mathbb{A}(\boldsymbol{\alpha}^{(1)}) - \mathbb{A}(\boldsymbol{\alpha}^{(1)})\right) \boldsymbol{u}^{(2)}, \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)} \right\rangle_n \\ &\qquad + \left\langle \left(\boldsymbol{\omega}(\boldsymbol{\alpha}^{(1)}) - \boldsymbol{\omega}(\boldsymbol{\alpha}^{(2)})\right) \bullet \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{w}^{(2)}_\tau| - |\boldsymbol{w}^{(1)}_\tau| \right\rangle_p \\ &\qquad + \left\langle \boldsymbol{L}(\boldsymbol{\alpha}^{(1)}) - \boldsymbol{L}(\boldsymbol{\alpha}^{(2)}), \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)} \right\rangle_n . \end{aligned}$$

Making use of (1.23), Lipschitz continuity of $\mathbb{A}$, $\boldsymbol{L}$, $\boldsymbol{\omega}$ and boundedness of $\mathfrak{F}$, we arrive at the following estimate:

$$\gamma\|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\|^2_n \leq c\|\boldsymbol{\alpha}^{(1)} - \boldsymbol{\alpha}^{(2)}\|_p \|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\|_n,$$

where the constant $c > 0$ does not depend on $\boldsymbol{\varphi}$ and $\boldsymbol{\alpha}^{(i)}$, $i = 1, 2$. From this and the definition of $\boldsymbol{a}^{(i)}$ then one obtains

$$\begin{aligned} \|\boldsymbol{u}^{(1)} - \boldsymbol{u}^{(2)}\|_n &\leq \|\boldsymbol{a}^{(1)} - \boldsymbol{a}^{(2)}\|_n + \|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\|_n \\ &\leq (1 + c)\|\boldsymbol{\alpha}^{(1)} - \boldsymbol{\alpha}^{(2)}\|_p. \end{aligned} \qquad (2.15)$$

To estimate the difference of the Lagrange multipliers $\|\boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)}\|_p$ we proceed as in the proof of the previous lemma. In particular, from the first inequality of $(\bar{A}(\boldsymbol{\alpha}^{(i)},\boldsymbol{\varphi}))$ we have

$$\left\langle \mathbb{A}(\boldsymbol{\alpha}^{(i)})\boldsymbol{u}^{(i)},\boldsymbol{v}\right\rangle_n = \left\langle \boldsymbol{L}(\boldsymbol{\alpha}^{(i)}),\boldsymbol{v}\right\rangle_n + \left\langle \boldsymbol{\lambda}^{(i)},\boldsymbol{v}_\nu\right\rangle_p \quad \forall \boldsymbol{v}\in\mathbb{R}^n,\ \boldsymbol{v}_\tau=\boldsymbol{0},\ i=1,2.$$

Subtracting the two equations for $i=1,2$ yields

$$\begin{aligned}\left\langle \boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)},\boldsymbol{v}_\nu\right\rangle_p =&\left\langle \left(\mathbb{A}(\boldsymbol{\alpha}^{(1)})-\mathbb{A}(\boldsymbol{\alpha}^{(2)})\right)\boldsymbol{u}^{(2)},\boldsymbol{v}\right\rangle_n + \left\langle \mathbb{A}(\boldsymbol{\alpha}^{(2)})\left(\boldsymbol{u}^{(1)}-\boldsymbol{u}^{(2)}\right),\boldsymbol{v}\right\rangle_n \\ &+ \left\langle \boldsymbol{L}(\boldsymbol{\alpha}^{(1)})-\boldsymbol{L}(\boldsymbol{\alpha}^{(2)}),\boldsymbol{v}\right\rangle_n \quad \forall \boldsymbol{v}\in\mathbb{R}^n, \boldsymbol{v}_\tau=\boldsymbol{0}.\end{aligned}$$

Proceeding exactly as in the proof of the previous lemma, i.e., divide by $\|\boldsymbol{v}\|_n$ and take supremum over $\mathcal{S}:=\{\boldsymbol{v}\in\mathbb{R}^n \mid \boldsymbol{v}_\nu\neq\boldsymbol{0},\ \text{the remaining components of } \boldsymbol{v} \text{ are } 0\}$, we arrive at

$$\|\boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)}\|_p = \sup_{\boldsymbol{v}\in\mathcal{S}} \frac{\left\langle \boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)},\boldsymbol{v}_\nu\right\rangle_p}{\|\boldsymbol{v}\|_n} \leq c\|\boldsymbol{\alpha}^{(1)}-\boldsymbol{\alpha}^{(2)}\|_p. \tag{2.16}$$

Here we made use of the Lipschitz continuity of $\mathbb{A}$ and $\boldsymbol{L}$, Lemma 2(i), as well as (2.15). Finally, the proof is finished by adding (2.15) and (2.16). ☐

The main result now reads as follows.

Theorem 16. *Let $\mathfrak{F}$ be Lipschitz with a sufficiently small modulus $C_{lip}>0$ so that Lemma 4 holds with $q<1$. Then $\widetilde{S}$, defined in* (2.11), *is single-valued and Lipschitz in $\widetilde{U}_{ad}$.*

Proof. Comparing the constants q and $\bar{K}$ from (2.14) and (1.39), respectively, one easily finds that $q<1$ implies that the assumption of Theorem 10 is satisfied. Thus $(M(\boldsymbol{\alpha}))$ are uniquely solvable for each $\boldsymbol{\alpha}\in\widetilde{U}_{ad}$ and hence $\widetilde{S}$ is single-valued.

Now, let $\boldsymbol{\alpha},\bar{\boldsymbol{\alpha}}\in\widetilde{U}_{ad}$ be given and denote the solutions to $(M(\boldsymbol{\alpha}))$ and $(M(\bar{\boldsymbol{\alpha}}))$ by $(\boldsymbol{u},\boldsymbol{\lambda})\in\mathbb{R}^n\times\mathbb{R}^p_+$ and $(\bar{\boldsymbol{u}},\bar{\boldsymbol{\lambda}})\in\mathbb{R}^n\times\mathbb{R}^p_+$, respectively. Since the corresponding mappings Ψ_α and $\Psi_{\bar{\alpha}}$ (cf. (1.36)) are contractive, these solutions may be revealed by the method of successive approximations in the following way.

Choose an arbitrary $\boldsymbol{\varphi}^{(0)}\in\mathbb{R}^p_+$ and compute the solutions to $(\bar{A}(\boldsymbol{\alpha},\boldsymbol{\varphi}^{(0)}))$ and $(\bar{A}(\bar{\boldsymbol{\alpha}},\boldsymbol{\varphi}^{(0)}))$—let us denote them by $(\boldsymbol{u}^{(0)},\boldsymbol{\lambda}^{(0)})\in\mathbb{R}^n\times\mathbb{R}^p_+$ and $(\bar{\boldsymbol{u}}^{(0)},\bar{\boldsymbol{\lambda}}^{(0)})\in\mathbb{R}^n\times\mathbb{R}^p_+$. Then set $\boldsymbol{\varphi}^{(1)}:=\Psi_\alpha(\boldsymbol{\varphi}^{(0)})$ and $\bar{\boldsymbol{\varphi}}^{(1)}:=\Psi_{\bar{\alpha}}(\boldsymbol{\varphi}^{(0)})$. By Lemma 5 we readily know that

$$\|(\boldsymbol{u}^{(0)},\boldsymbol{\lambda}^{(0)})-(\bar{\boldsymbol{u}}^{(0)},\bar{\boldsymbol{\lambda}}^{(0)})\|_{n+p}\leq c\|\boldsymbol{\alpha}-\bar{\boldsymbol{\alpha}}\|_p, \tag{2.17}$$

and hence also

$$\|\boldsymbol{\varphi}^{(1)}-\bar{\boldsymbol{\varphi}}^{(1)}\|_p = \left\||\boldsymbol{u}^{(0)}_\tau|-|\bar{\boldsymbol{u}}^{(0)}_\tau|\right\|_n \leq \|\boldsymbol{u}^{(0)}-\bar{\boldsymbol{u}}^{(0)}\|_n \leq c\|\boldsymbol{\alpha}-\bar{\boldsymbol{\alpha}}\|_p. \tag{2.18}$$

Now, solve the problems $(\bar{A}(\boldsymbol{\alpha},\boldsymbol{\varphi}^{(1)}))$ and $(\bar{A}(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{\varphi}}^{(1)}))$ to obtain $(\boldsymbol{u}^{(1)},\boldsymbol{\lambda}^{(1)})\in\mathbb{R}^n\times\mathbb{R}^p_+$ and $(\bar{\boldsymbol{u}}^{(1)},\bar{\boldsymbol{\lambda}}^{(1)})\in\mathbb{R}^n\times\mathbb{R}^p_+$. Further, denote the solution to $(\bar{A}(\boldsymbol{\alpha},\bar{\boldsymbol{\varphi}}^{(1)}))$ by $(\boldsymbol{U}^{(1)},\boldsymbol{\Lambda}^{(1)})\in\mathbb{R}^n\times\mathbb{R}^p_+$. Thus, we may estimate:

$$\begin{aligned}\|(\boldsymbol{u}^{(1)},\boldsymbol{\lambda}^{(1)})-(\bar{\boldsymbol{u}}^{(1)},\bar{\boldsymbol{\lambda}}^{(1)})\|_{n+p} &\leq \|(\boldsymbol{u}^{(1)},\boldsymbol{\lambda}^{(1)})-(\boldsymbol{U}^{(1)},\boldsymbol{\Lambda}^{(1)})\|_{n+p} \\ &\quad + \|(\boldsymbol{U}^{(1)},\boldsymbol{\Lambda}^{(1)})-(\bar{\boldsymbol{u}}^{(1)},\bar{\boldsymbol{\lambda}}^{(1)})\|_{n+p} \\ &\leq q\|\boldsymbol{\varphi}^{(1)}-\bar{\boldsymbol{\varphi}}^{(1)}\|_p + c\|\boldsymbol{\alpha}-\bar{\boldsymbol{\alpha}}\|_p \\ &\leq c(1+q)\|\boldsymbol{\alpha}-\bar{\boldsymbol{\alpha}}\|_p,\end{aligned}$$

as follows from Lemmas 4, 5 and (2.18). Continuing this iterative process, in the kth step one has $(\boldsymbol{u}^{(k)}, \boldsymbol{\lambda}^{(k)})$ and $(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)})$, the solutions to $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^{(k)}))$ and $(\bar{A}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\varphi}}^{(k)}))$, respectively, along with the estimate:

$$\begin{aligned}\|(\boldsymbol{u}^{(k)}, \boldsymbol{\lambda}^{(k)}) - (\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}}^{(k)})\|_{n+p} &\leq c(1 + q + q^2 + \cdots + q^k)\|\boldsymbol{\alpha} - \bar{\boldsymbol{\alpha}}\|_p \\ &\leq \frac{c}{1-q}\|\boldsymbol{\alpha} - \bar{\boldsymbol{\alpha}}\|_p, \end{aligned} \tag{2.19}$$

since $q < 1$ by assumption. Then, one sets $\boldsymbol{\varphi}^{(k+1)} := \Psi_\alpha(\boldsymbol{\varphi}^{(k)})$, $\bar{\boldsymbol{\varphi}}^{(k+1)} := \Psi_{\bar{\alpha}}(\bar{\boldsymbol{\varphi}}^{(k)})$, and starts the iteration loop with $k := k + 1$.

The sequences $\{\boldsymbol{\varphi}^{(k)}\}$ and $\{\bar{\boldsymbol{\varphi}}^{(k)}\}$ generated by this process converge to the unique fixed points of the mappings Ψ_α and $\Psi_{\bar{\alpha}}$, resp.; the sequences $\{(\boldsymbol{u}^{(k)}, \boldsymbol{\lambda}^{(k)})\}$, $\{(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)})\}$ converge to the (unique) solutions of $(M(\boldsymbol{\alpha}))$ and $(M(\bar{\boldsymbol{\alpha}}))$, resp. Thus it is sufficient to pass to the limit as $k \to \infty$ in (2.19) to obtain the assertion of the theorem. □

2.3.2 Stability with respect to the load vector

In addition to Theorem 16 we shall need another stability result, namely the one with respect to the load vector $\boldsymbol{L}$. Since $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$ will be fixed and $\boldsymbol{L} \in \mathbb{R}^n$ the parameter, we adjust the notation to reflect this fact and write $(\bar{A}(\boldsymbol{L}, \boldsymbol{\varphi}))$, $(M(\boldsymbol{L}))$, Ψ_L instead of $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$, $(M(\boldsymbol{\alpha}))$, Ψ_α, etc.

Lemma 6. *Let $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$ be fixed and the assumptions of Theorem 16 hold true. Then there exists a constant $c > 0$, not depending on $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$, such that*

$$\|(\boldsymbol{u}, \boldsymbol{\lambda}) - (\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}})\|_{n+p} \leq c\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n \quad \forall \boldsymbol{L}, \bar{\boldsymbol{L}} \in \mathbb{R}^n,$$

where $(\boldsymbol{u}, \boldsymbol{\lambda}), (\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ denote the (unique) solutions of $(M(\boldsymbol{L}))$, $(M(\bar{\boldsymbol{L}}))$, respectively.

Proof. We merely sketch the proof, since it employs the same fixed-point technique as the proof of Theorem 16.

Let $\boldsymbol{\varphi}^{(0)} \in \mathbb{R}^p_+$ be arbitrary and denote the solutions to $(\bar{A}(\boldsymbol{L}, \boldsymbol{\varphi}^{(0)}))$, $(\bar{A}(\bar{\boldsymbol{L}}, \boldsymbol{\varphi}^{(0)}))$ by $(\boldsymbol{u}^{(0)}, \boldsymbol{\lambda}^{(0)}), (\bar{\boldsymbol{u}}^{(0)}, \bar{\boldsymbol{\lambda}}^{(0)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$, respectively. Then by Lemma 2(ii) we know that there exists a $c > 0$, independent of $\boldsymbol{\alpha}$, $\boldsymbol{L}$, $\bar{\boldsymbol{L}}$ and $\boldsymbol{\varphi}^{(0)}$, such that

$$\|(\boldsymbol{u}^{(0)}, \boldsymbol{\lambda}^{(0)}) - (\bar{\boldsymbol{u}}^{(0)}, \bar{\boldsymbol{\lambda}}^{(0)})\|_{n+p} \leq c\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n. \tag{2.20}$$

Next, we define $\boldsymbol{\varphi}^{(1)} := \Psi_L(\boldsymbol{\varphi}^{(0)})$ and $\bar{\boldsymbol{\varphi}}^{(1)} := \Psi_{\bar{L}}(\boldsymbol{\varphi}^{(0)})$. The respective solutions to $(\bar{A}(\boldsymbol{L}, \boldsymbol{\varphi}^{(1)}))$ and $(\bar{A}(\bar{\boldsymbol{L}}, \bar{\boldsymbol{\varphi}}^{(1)}))$ shall be denoted by $(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}), (\bar{\boldsymbol{u}}^{(1)}, \bar{\boldsymbol{\lambda}}^{(1)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$. In addition, we solve the problem $(\bar{A}(\boldsymbol{L}, \bar{\boldsymbol{\varphi}}^{(1)}))$ and signify its solution by $(\boldsymbol{U}^{(1)}, \boldsymbol{\Lambda}^{(1)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$. Then, due to Lemma 4 and (2.20), one has:

$$\begin{aligned}\|(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}) - (\boldsymbol{U}^{(1)}, \boldsymbol{\Lambda}^{(1)})\|_{n+p} &\leq q\|\boldsymbol{\varphi}^{(1)} - \bar{\boldsymbol{\varphi}}^{(1)}\|_p \\ &\leq q\|\boldsymbol{u}^{(0)} - \bar{\boldsymbol{u}}^{(0)}\|_n \\ &\leq qc\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n. \end{aligned} \tag{2.21}$$

On the other hand, from Lemma 2(ii) it immediately follows that

$$\|(\boldsymbol{U}^{(1)}, \boldsymbol{\Lambda}^{(1)}) - (\bar{\boldsymbol{u}}^{(1)}, \bar{\boldsymbol{\lambda}}^{(1)})\|_{n+p} \leq c\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n. \tag{2.22}$$

Adding (2.21), (2.22), and using the triangle inequality, we get

$$\|(\boldsymbol{u}^{(1)}, \boldsymbol{\lambda}^{(1)}) - (\bar{\boldsymbol{u}}^{(1)}, \bar{\boldsymbol{\lambda}}^{(1)})\|_{n+p} \leq c(1+q)\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n. \tag{2.23}$$

The rest of the proof may be conducted in the same manner as was done in Theorem 16. Indeed, one defines the sequences $\{\boldsymbol{\varphi}^{(k)}\}, \{\bar{\boldsymbol{\varphi}}^{(k)}\} \subset \mathbb{R}^p_+$ by $\boldsymbol{\varphi}^{(k)} := \Psi_L(\boldsymbol{\varphi}^{(k-1)})$ and $\bar{\boldsymbol{\varphi}}^{(k)} := \Psi_{\bar{L}}(\bar{\boldsymbol{\varphi}}^{(k-1)})$, $k = 2, 3, \ldots$, resp.; the elements of the sequences $\{(\boldsymbol{u}^{(k)}, \boldsymbol{\lambda}^{(k)})\}$, $\{(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)})\}$, $\{(\boldsymbol{U}^{(k)}, \boldsymbol{\Lambda}^{(k)})\} \subset \mathbb{R}^n \times \mathbb{R}^p_+$ are then defined, in this order, as the unique solutions of problems $(\bar{A}(\boldsymbol{L}, \boldsymbol{\varphi}^{(k)}))$, $(\bar{A}(\bar{\boldsymbol{L}}, \bar{\boldsymbol{\varphi}}^{(k)}))$, $(\bar{A}(\boldsymbol{L}, \bar{\boldsymbol{\varphi}}^{(k)}))$ $\forall k \in \mathbb{N}, k \geq 2$. By induction one may prove the estimate (cf. (2.19) and (2.23))

$$\|(\boldsymbol{u}^{(k)}, \boldsymbol{\lambda}^{(k)}) - (\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)})\|_{n+p} \leq \frac{c}{1-q}\|\boldsymbol{L} - \bar{\boldsymbol{L}}\|_n, \tag{2.24}$$

since by assumption $q < 1$ holds. The desired result is then obtained by taking limit in (2.24) as $k \to \infty$; for details we kindly refer to the proof of the previous theorem. ☐

2.4 Implicit Programming

Having Theorem 16 at hand, we return to the shape optimization problem $(\mathbb{P})$ and present a method for its solution. To this end, let us assume, that the assumptions of Theorem 16 are satified, i.e., $\mathfrak{F}$ is Lipschitz with a sufficiently small modulus. In addition, let the cost functional J be *continuously differentiable.* In fact, this smoothness assumption is superfluous and is imposed only for the sake of simplicity. As it will become apparent, a locally Lipschitz J would work—in theory—just as fine.

Provided the assumptions above are satisfied and $\widetilde{S}$ is single-valued, we may apply the Implicit Programming (ImP) approach to the solution of $(\mathbb{P})$. This consists of reformulating $(\mathbb{P})$ as the following nonlinear program (NLP):

$$\left.\begin{array}{ll} \text{minimize} & \mathcal{J}(\boldsymbol{\alpha}) := J(\boldsymbol{\alpha}, \widetilde{S}(\boldsymbol{\alpha})), \\ \text{subj. to} & \boldsymbol{\alpha} \in \widetilde{U}_{ad}. \end{array}\right\} \tag{$\widetilde{\mathbb{P}}$}$$

Eliminating the equilibrium constraint, only the geometric constraint $\boldsymbol{\alpha} \in \widetilde{U}_{ad}$ remains, in which the compact, convex feasible set is given by simple box-constraints and linear inequality (and, if $C_{21} = C_{22}$, also equality) constraints (cf. (2.5)).

2.4.1 Selecting a minimization algorithm

There are several aspects, that have to be taken into account when attempting to solve $(\widetilde{\mathbb{P}})$:

- Although J was assumed continuously differentiable (or even smoother), due to the Lipschitz continuity of $\widetilde{S}$, the reduced cost functional $\mathcal{J} : \widetilde{U}_{ad} \to \mathbb{R}$ is only Lipschitz, in general.

- $\widetilde{S}$ is not convex, in general; therefore, the same applies to $\mathcal{J}$ as well.

- Each function evaluation of $\mathcal{J}$ is costly, since—by means of $\widetilde{S}$—it involves solving a frictional contact problem, where the coefficient of friction depends on the solution.

- Typically, in practical computations, the optimized boundary segment is not parametrized by (nonsmooth) piecewise linear functions, but smooth curves, like piecewise quadratic or cubic Bézier or spline functions. Besides being smooth, they can be controlled by relatively few points to achieve satisfactory design. Therefore, in most cases, one may assume that the dimension of $\widetilde{U}_{ad}$ is at most "moderately" large (in the sense of [23]).

From the first two points it follows that $(\widetilde{\mathbb{P}})$ has to be solved by a method of nonsmooth and nonconvex optimization, whereas the third point basically rules out derivative-free methods—they typically require orders of magnitude more function evaluations, than algorithms based on first (and second) order (sub)gradients.

As no additional structural property, that could be exploited by the minimization algorithm, is known a priori, we opt for the Bundle Trust [52, 55] and Proximal Bundle [39, 40] methods. In general, bundle methods have turned out to be the method of choice for the solution of small to medium scale, nonsmooth, nonconvex optimization problems, without extra knowledge about their structure—see the comparison in [23] and also [38].

2.4.2 Computing a subgradient

Bundle methods are iterative methods for minimizing the locally Lipschitz objective function $\mathcal{J}$, that require at each step $\bar{\boldsymbol{\alpha}} \in \widetilde{U}_{ad}$:

(i) the function value $\mathcal{J}(\bar{\boldsymbol{\alpha}})$, and

(ii) one arbitrary subgradient $\boldsymbol{\xi} \in \bar{\partial}\mathcal{J}(\bar{\boldsymbol{\alpha}})$ from the Clarke subdifferential [7].

As readily seen, in order to provide $\mathcal{J}(\bar{\boldsymbol{\alpha}})$, one has to evaluate $\bar{\boldsymbol{y}} := \widetilde{S}(\bar{\boldsymbol{\alpha}})$, i.e., it is necessary to solve a Signorini problem with Tresca friction and a solution-dependent coefficient of friction. Assume, we are able to solve $(GE(\bar{\boldsymbol{\alpha}}))$ and let us focus on task (ii). By the chain rule [7, Theorem 2.6.6] we have:

$$\bar{\partial}\mathcal{J}(\bar{\boldsymbol{\alpha}}) = \nabla_\alpha J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \big(\bar{\partial}\widetilde{S}(\bar{\boldsymbol{\alpha}})\big)^T \nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}). \tag{2.25}$$

This means that determining an element of (2.25) involves computing a generalized Jacobian of the nonsmooth, implicitly defined control-to-state mapping $\widetilde{S}$. This can be conducted essentially in two different ways:

(j) If $\widetilde{S}$ happens to be piecewise C^1 (PC^1), it is convenient to obtain the desired subgradient completely within the generalized differential calculus of Clarke (specialized implicit function theorems are provided for example in [48]). This way has been applied, e.g., in [3].

(jj) If the PC^1 nature of the control-to-state mapping cannot be guaranteed, it seems reasonable to perform sensitivity analysis via the generalized differential calculus of Mordukhovich [36] which is richer concerning specialized calculus rules. The paper [4] may serve as an example for the viability of this approach.

Since in the considered model we have to do with rather complicated nonsmooth and set-valued mappings, we have chosen the second approach. In the next section it is shown how to compute an approximation of a Clarke subgradient from the set (2.25) by means of the generalized differential calculus of B. Mordukhovich.

2.5 Sensitivity analysis

Let $\bar{\boldsymbol{\alpha}} \in \widetilde{U}_{ad}$ be arbitrary and denote $\bar{\boldsymbol{y}} := \widetilde{S}(\bar{\boldsymbol{\alpha}})$. We start with the following fact, providing a link between the differential operators of interest from the Clarke and Mordukhovich calculus.

Lemma 7. *For any $\boldsymbol{y}^* \in \mathbb{R}^{3p}$ it holds that $D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*) \neq \emptyset$ and*

$$\big(\bar{\partial}\widetilde{S}(\bar{\boldsymbol{\alpha}})\big)^T \boldsymbol{y}^* = \operatorname{conv} D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*). \tag{2.26}$$

Proof. Follows from the Lipschitz continuity of $\widetilde{S}$ and formula (2.23) in [35]. ☐

Comparing (2.26) and (2.25) we see that for our purposes it is sufficient to compute one $\boldsymbol{p}^* \in D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$; then, setting

$$\boldsymbol{\xi} := \nabla_\alpha J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \boldsymbol{p}^* \tag{2.27}$$

we are done. However, this is not straightforward, since $\widetilde{S}$ is defined via an implicit relation. In order to express its coderivative $D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})$ in terms of $\boldsymbol{F}$ and $\widetilde{Q}$, we start with the following observation:

$$\begin{aligned}
\operatorname{Gr}\widetilde{S} &= \{(\boldsymbol{\alpha}, \boldsymbol{y}) \in \widetilde{U}_{ad} \times \mathbb{R}^{3p} \mid -\boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) \in \widetilde{Q}(\boldsymbol{\alpha}, \boldsymbol{y})\} \\
&= \{(\boldsymbol{\alpha}, \boldsymbol{y}) \in \widetilde{U}_{ad} \times \mathbb{R}^{3p} \mid \Phi(\boldsymbol{\alpha}, \boldsymbol{y}) := (\boldsymbol{\alpha}, \boldsymbol{y}, -\boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y})) \in \operatorname{Gr}\widetilde{Q}\} \qquad (2.28) \\
&= \Phi^{-1}\big(\operatorname{Gr}\widetilde{Q}\big). \qquad (2.29)
\end{aligned}$$

To be able to compute normal cones to the set (2.29), one has to verify a calmness condition, as presented below.

Lemma 8. *Let $\bar{\boldsymbol{\alpha}} \in \widetilde{U}_{ad}$ be fixed, $\bar{\boldsymbol{y}} := \widetilde{S}(\bar{\boldsymbol{\alpha}})$ and the mapping $\Phi : \widetilde{U}_{ad} \times \mathbb{R}^{3p} \to \widetilde{U}_{ad} \times \mathbb{R}^{3p} \times \mathbb{R}^{3p}$ be defined by (2.28). Then the multifunction $M : \mathbb{R}^p \times \mathbb{R}^{3p} \times \mathbb{R}^{3p} \rightrightarrows \widetilde{U}_{ad} \times \mathbb{R}^{3p}$ given by*

$$M : \boldsymbol{p} \mapsto \{(\boldsymbol{\alpha}, \boldsymbol{y}) \mid \boldsymbol{p} + \Phi(\boldsymbol{\alpha}, \boldsymbol{y}) \in \operatorname{Gr}\widetilde{Q}\}$$

is calm at $(\mathbf{0}, \mathbf{0}, \mathbf{0}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$.

Proof. If M was not calm at $(\mathbf{0}, \mathbf{0}, \mathbf{0}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$, one could easily disprove calmness of the following multifunction $\tilde{M} : \mathbb{R}^{3p} \rightrightarrows \widetilde{U}_{ad} \times \mathbb{R}^{3p}$ at $(\mathbf{0}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$:

$$\tilde{M} : \tilde{\boldsymbol{p}} \mapsto \{(\boldsymbol{\alpha}, \boldsymbol{y}) \mid (\mathbf{0}, \mathbf{0}, \tilde{\boldsymbol{p}}) + \Phi(\boldsymbol{\alpha}, \boldsymbol{y}) \in \operatorname{Gr}\widetilde{Q}\}.$$

Indeed, suppose that there exist sequences $\boldsymbol{p}^{(i)} = (\boldsymbol{p}_1^{(i)}, \boldsymbol{p}_2^{(i)}, \boldsymbol{p}_3^{(i)}) \to (\mathbf{0}, \mathbf{0}, \mathbf{0}) \in \mathbb{R}^p \times \mathbb{R}^{3p} \times \mathbb{R}^{3p}$ and $(\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}) \to (\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$, $(\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}) \in M(\boldsymbol{p}^{(i)})$ such that

$$\operatorname{dist}((\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}), M(\mathbf{0}, \mathbf{0}, \mathbf{0})) \geq i\|\boldsymbol{p}^{(i)}\|_{p+3p+3p} \quad \forall i \in \mathbb{N}. \tag{2.30}$$

Let us put $(\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}) := (\boldsymbol{\alpha}^{(i)} + \boldsymbol{p}_1^{(i)}, \boldsymbol{y}^{(i)} + \boldsymbol{p}_2^{(i)})$ so that the relation $(\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}) \in M(\boldsymbol{p}^{(i)})$ can be rewritten as

$$\tilde{\boldsymbol{p}}^{(i)} \in \boldsymbol{F}(\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}) + \widetilde{Q}(\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}), \tag{2.31}$$

with

$$\tilde{\boldsymbol{p}}^{(i)} := \boldsymbol{p}_3^{(i)} - \boldsymbol{F}(\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}) - \boldsymbol{F}(\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}). \tag{2.32}$$

Since $\boldsymbol{F}$ is locally Lipschitz, one has from (2.31), (2.32):

$$\begin{aligned} &\|\tilde{\boldsymbol{p}}^{(i)}\|_{3p} \le c\|\boldsymbol{p}^{(i)}\|_{p+3p+3p} \quad \text{and} \quad \tilde{\boldsymbol{p}}^{(i)} \to \boldsymbol{0} \in \mathbb{R}^{3p}, \\ &(\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}) \in \tilde{M}(\tilde{\boldsymbol{p}}^{(i)}), \quad \text{and} \quad (\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}) \to (\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}). \end{aligned}$$

Thus we can estimate:

$$\begin{aligned} \operatorname{dist}((\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}), \tilde{M}(\boldsymbol{0})) &= \operatorname{dist}((\boldsymbol{\beta}^{(i)}, \boldsymbol{z}^{(i)}), M(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})) \\ &\ge \operatorname{dist}((\boldsymbol{\alpha}^{(i)}, \boldsymbol{y}^{(i)}), M(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})) - \|\boldsymbol{p}^{(i)}\|_{p+3p+3p} \ge (i-1)\|\boldsymbol{p}^{(i)}\|_{p+3p+3p} \\ &\ge \frac{i-1}{c}\|\tilde{\boldsymbol{p}}^{(i)}\|_{3p} \end{aligned}$$

and the claim has been verified.

Therefore it is sufficient to show that $\tilde{M}$ is calm at $(\boldsymbol{0}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$. To this end, let $\tilde{\boldsymbol{p}} \in \mathbb{R}^{3p}$ be given. Then

$$(\boldsymbol{\alpha}, \boldsymbol{y}) \in \tilde{M}(\tilde{\boldsymbol{p}}) \quad \Leftrightarrow \quad \tilde{\boldsymbol{p}} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + \widetilde{Q}(\boldsymbol{\alpha}, \boldsymbol{y}),$$

i.e., written componentwise for $\boldsymbol{y} = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda})$ and $\tilde{\boldsymbol{p}} = (\tilde{\boldsymbol{p}}_1, \tilde{\boldsymbol{p}}_2, \tilde{\boldsymbol{p}}_3) \in \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p$:

$$\left.\begin{aligned} &\tilde{\boldsymbol{p}}_1 \in \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + \widetilde{Q}_\tau(\boldsymbol{\alpha}, \boldsymbol{u}_\tau) \\ &\tilde{\boldsymbol{p}}_2 = \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\boldsymbol{\alpha}) \\ &\tilde{\boldsymbol{p}}_3 \in \boldsymbol{u}_\nu + \boldsymbol{\alpha} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}). \end{aligned}\right\} \tag{2.33}$$

Introducing the new variable $\tilde{\boldsymbol{y}} := (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu - \tilde{\boldsymbol{p}}_3, \boldsymbol{\lambda})$, we see that $(\boldsymbol{\alpha}, \tilde{\boldsymbol{y}})$ solves (2.6) with the load vector

$$\tilde{\boldsymbol{l}} := \begin{bmatrix} \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + \tilde{\boldsymbol{p}}_1 - \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\tilde{\boldsymbol{p}}_3 \\ \boldsymbol{L}_\nu(\boldsymbol{\alpha}) + \tilde{\boldsymbol{p}}_2 - \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\tilde{\boldsymbol{p}}_3 \\ -\boldsymbol{\alpha} \end{bmatrix}.$$

From Theorem 10 it follows that $(\boldsymbol{\alpha}, \tilde{\boldsymbol{y}})$ is the only solution to the perturbed GE (2.33). Denoting $(\boldsymbol{\alpha}, \boldsymbol{y}^*) \in \operatorname{Gr} \widetilde{S} = \tilde{M}(\boldsymbol{0})$ the solution to (2.6) with the original load vector $\boldsymbol{l} = [\boldsymbol{L}_\tau(\boldsymbol{\alpha}), \boldsymbol{L}_\nu(\boldsymbol{\alpha}), -\boldsymbol{\alpha}]^T$, we obtain from the triangle inequality and Lemma 6:

$$\begin{aligned} \|(\boldsymbol{\alpha}, \boldsymbol{y}) - (\boldsymbol{\alpha}, \boldsymbol{y}^*)\|_{p+3p} &\le \|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|_{3p} + \|\tilde{\boldsymbol{y}} - \boldsymbol{y}^*\|_{3p} \\ &\le \|\tilde{\boldsymbol{p}}_3\|_p + c\|\tilde{\boldsymbol{l}} - \boldsymbol{l}\|_{3p} \\ &\le c\|\tilde{\boldsymbol{p}}\|_{3p}, \end{aligned}$$

where $c > 0$ does not depend on $\boldsymbol{\alpha}$. From this the required calmness property follows easily. □

2.5.1 The adjoint GE

The following result (see [27, Theorem 2]) facilitates the computation of the adjoint variable $\boldsymbol{p}^* \in D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*)$, needed in (2.27).

Theorem 17. *Consider a reference pair* $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} \widetilde{S}$ *and let* $\boldsymbol{y}^* \in \mathbb{R}^{3p}$ *be arbitrary.*

(i) Let $(\boldsymbol{p}^*, \boldsymbol{v}^*) \in \mathbb{R}^p \times \mathbb{R}^{3p}$ *be a solution to the regular adjoint GE:*

$$\begin{bmatrix} \boldsymbol{p}^* \\ -\boldsymbol{y}^* \end{bmatrix} \in \nabla \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^* + \widehat{D}^* \widetilde{Q}(\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{v}^*). \tag{RAGE}$$

Then $\boldsymbol{p}^* \in D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*)$.

(ii) For every $\boldsymbol{p}^* \in D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*)$ *there exists a vector* $\boldsymbol{v}^* \in \mathbb{R}^{3p}$ *such that* $(\boldsymbol{p}^*, \boldsymbol{v}^*)$ *is a solution of the (limiting) adjoint GE:*

$$\begin{bmatrix} \boldsymbol{p}^* \\ -\boldsymbol{y}^* \end{bmatrix} \in \nabla \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^* + D^* \widetilde{Q}(\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{v}^*). \tag{AGE}$$

Proof. The first assertion follows immediately from [47, Theorem 10.6]. The second one is implied by [21, Theorem 4.1], whose assumptions are fulfilled by virtue of Lemma 8. □

Note that due to Lipschitz continuity of $\widetilde{S}$, (AGE) attains at least one solution $\boldsymbol{p}^*$ (cf. Lemma 7) and whenever $\widetilde{Q}$ is normally regular at $\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$, i.e., $\widehat{N}_{\operatorname{Gr} \widetilde{Q}}(\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})) = N_{\operatorname{Gr} \widetilde{Q}}(\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$, (RAGE) and (AGE) coincide. In this case

$$D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\boldsymbol{y}^*) = \{\boldsymbol{p}^* \in \mathbb{R}^p \mid \exists \boldsymbol{v}^* \in \mathbb{R}^{3p} \text{ such that } (\boldsymbol{p}^*, \boldsymbol{v}^*) \text{ solves (AGE)}\}.$$

On the other hand, in the nonregular case (RAGE) may be difficult to solve or not solvable at all. Therefore the computation of the desired subgradient $\boldsymbol{\xi} \in \overline{\partial} \mathcal{J}(\bar{\boldsymbol{\alpha}})$ is usually done via the (AGE), while accepting the fact that at nonregular points the computed vector may lie outside of $\overline{\partial} \mathcal{J}(\bar{\boldsymbol{\alpha}})$. In such cases the employed optimization algorithm might collapse and $\boldsymbol{\xi}$ has to be replaced by a correct subgradient.

In light of the previous paragraph we will focus on the solution of the (AGE). In particular, we shall express the most difficult part of (AGE), the coderivative $D^*\widetilde{Q}(\Phi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$ in terms of the problem data.

Computation of $D^*\widetilde{Q}$

First of all, note that the components of $\widetilde{Q}$ are *decoupled*–the first component $\widetilde{Q}_\tau$ depends on $\boldsymbol{\alpha}$ and $\boldsymbol{u}_\tau$, whereas the third component computes the normal cone to $\mathbb{R}^p_+$ only at $\boldsymbol{\lambda}$. Actually, this fact is a consequence of the assumed model of given friction, since $\widetilde{Q}_\tau$ reflects the friction condition and the third component corresponds to the nonpenetration condition. This way, the coderivative of $\widetilde{Q}$ may be computed componentwise [47, Example 6.10]:

$$\forall \boldsymbol{q}^* \in \mathbb{R}^{3p}: \quad D^*\widetilde{Q}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{q}})(\boldsymbol{q}^*) = \begin{bmatrix} D^*\widetilde{Q}_\tau(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}_1, \bar{\boldsymbol{q}}_1)(\boldsymbol{q}_1^*) \\ 0 \\ D^* N_{\mathbb{R}^p_+}(\bar{\boldsymbol{y}}_3, \bar{\boldsymbol{q}}_3)(\boldsymbol{q}_3^*) \end{bmatrix}, \tag{2.34}$$

at any reference point $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{q}}) \in \operatorname{Gr} \widetilde{Q}$, where $\bar{\boldsymbol{y}} = (\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{y}}_2, \bar{\boldsymbol{y}}_3)$, $\bar{\boldsymbol{q}} = (\bar{\boldsymbol{q}}_1, \bar{\boldsymbol{q}}_2, \bar{\boldsymbol{q}}_3)$, $\boldsymbol{q}^* = (\boldsymbol{q}_1^*, \boldsymbol{q}_2^*, \boldsymbol{q}_3^*) \in \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}^p$.

The formula for the coderivative of the normal cone mapping to $\mathbb{R}^p_+$ is well-known in the literature and may be very easily derived from the definition of the coderivative (see also Figure 2.1).

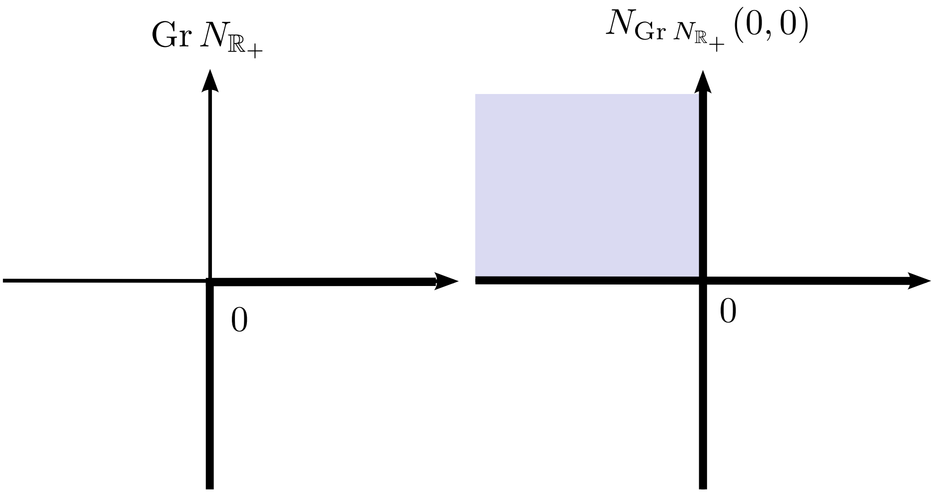


Figure 2.1: Graph of $N_{\mathbb{R}_+}$ and the normal cone to this set at $(0, 0)$.

Proposition 2. *Let* $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in \operatorname{Gr} N_{\mathbb{R}^p_+}$ *and* $\boldsymbol{b}^* \in \mathbb{R}^p$ *arbitrary. Then*

$$\boldsymbol{a}^* \in D^* N_{\mathbb{R}^p_+}(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) \quad \Leftrightarrow \quad a_i^* \in D^* N_{\mathbb{R}_+}(\bar{a}_i, \bar{b}_i)(b_i^*) \quad \forall i = 1, \ldots, p,$$

where

(i) if $\bar{a}_i > 0$, $\bar{b}_i = 0$, *then*

$$D^* N_{\mathbb{R}_+}(\bar{a}_i, \bar{b}_i)(b_i^*) = \{0\};$$

(ii) if $\bar{a}_i = 0$, $\bar{b}_i < 0$, *then*

$$D^* N_{\mathbb{R}_+}(\bar{a}_i, \bar{b}_i)(b_i^*) = \begin{cases} \mathbb{R} & \text{if } b_i^* = 0, \\ \emptyset & \text{otherwise}; \end{cases}$$

(iii) if $\bar{a}_i = 0$, $\bar{b}_i = 0$, *then*

$$D^* N_{\mathbb{R}_+}(\bar{a}_i, \bar{b}_i)(b_i^*) = \begin{cases} \{0\} & \text{if } b_i^* > 0, \\ \mathbb{R}_- & \text{if } b_i^* < 0, \\ \mathbb{R} & \text{if } b_i^* = 0. \end{cases}$$

Proof. See [41, Lemma 2.2]. □

Remark 7. Observe, that $D^* N_{\mathbb{R}^p_+}(\boldsymbol{a}, \boldsymbol{b}) = D^*(\partial \delta_{\mathbb{R}^p_+})(\boldsymbol{a}, \boldsymbol{b})$, which is the definition of the second-order subdifferential $\partial^2 \delta_{\mathbb{R}^p_+}(\boldsymbol{a}, \boldsymbol{b})$. This means, that the coderivative $D^* \widetilde{Q}$ in (AGE) provides second-order (sub)gradient information.

In order to deal with the first component, let us write the multifunction $\widetilde{Q}_\tau : \mathbb{R}^p \times \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ as a composition of an outer multifunction Z_τ and an inner single-valued, smooth mapping Ψ:

$$\widetilde{Q}_\tau(\boldsymbol{\alpha}, \boldsymbol{u}) = \begin{bmatrix} \omega_1(\boldsymbol{\alpha})\mathfrak{F}(|\boldsymbol{u}_1|)\partial|\boldsymbol{u}_1| \\ \omega_2(\boldsymbol{\alpha})\mathfrak{F}(|\boldsymbol{u}_2|)\partial|\boldsymbol{u}_2| \\ \vdots \\ \omega_p(\boldsymbol{\alpha})\mathfrak{F}(|\boldsymbol{u}_p|)\partial|\boldsymbol{u}_p| \end{bmatrix} = (Z_\tau \circ \Psi)(\boldsymbol{\alpha}, \boldsymbol{u}), \tag{2.35}$$

where

$$\Psi = (\Psi_1, \dots, \Psi_p) : \mathbb{R}^p \times \mathbb{R}^p \to ((0,\infty) \times \mathbb{R})^p, \qquad \Psi_j(\boldsymbol{\alpha}, \boldsymbol{u}) := \big(\omega_j(\boldsymbol{\alpha}), u_j\big),$$

and

$$Z_\tau : ((0,\infty) \times \mathbb{R})^p \rightrightarrows \mathbb{R}^p, \quad \boldsymbol{y} \mapsto \big(Z(\boldsymbol{y}_1), \dots, Z(\boldsymbol{y}_p)\big),$$

with

$$Z : (0,\infty) \times \mathbb{R} \rightrightarrows \mathbb{R}, \quad (x_1, x_2) \mapsto x_1\mathfrak{F}(|x_2|)\partial|x_2|.$$

Now the chain rule from [47, Theorem 10.40] allows us to compute the coderivative of the composite multifunction (2.35) as follows:

Proposition 3. *Let $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{q}}) \in \operatorname{Gr} \widetilde{Q}_\tau$ be such that the following condition holds:*

$$\big(\ker \nabla\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})^T\big) \cap D^*Z_\tau(\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), \bar{\boldsymbol{q}})(\boldsymbol{0}) = \{\boldsymbol{0}\}. \tag{2.36}$$

Then for every $\boldsymbol{q}^* = (q_1^*, \dots, q_p^*) \in \mathbb{R}^p$ one has

$$\begin{aligned} D^*\widetilde{Q}_\tau(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{q}})(\boldsymbol{q}^*) &\subset \nabla\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})^T D^*Z_\tau(\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), \bar{\boldsymbol{q}})(\boldsymbol{q}^*) \\ &= \nabla\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})^T \begin{bmatrix} D^*Z(\Psi_1(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), q_1)(q_1^*) \\ D^*Z(\Psi_2(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), q_2)(q_2^*) \\ \vdots \\ D^*Z(\Psi_p(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), q_p)(q_p^*) \end{bmatrix}. \end{aligned} \tag{2.37}$$

Observe that the assertion of Proposition 3 requires the validity of the qualification condition (2.36). We are going to show that (2.36) is satisfied at all points $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{q}}) \in \operatorname{Gr} \widetilde{Q}_\tau$ and hence the assertion of Proposition 3 holds automatically.

Remark 8. The right inclusion above becomes equality at those points $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{q}})$, for which the multifunction Z_τ is normally regular at $(\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), \bar{\boldsymbol{q}})$ or $\nabla\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})$ is surjective. In other cases, however, the formula on the right-hand side may provide a vector outside of $D^*\widetilde{Q}_\tau$.

Let us look more closely at the second option, i.e., what does it mean for $\nabla\Psi$ to be surjective at $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})$. Recalling the definition of Ψ, its Jacobian $\nabla\Psi \in \mathbb{R}^{2p\times 2p}$ can be written in the block-matrix form

$$\nabla\Psi = \begin{bmatrix} \mathbb{J}_{11} & \mathbb{J}_{12} \\ \vdots & \vdots \\ \mathbb{J}_{p1} & \mathbb{J}_{p2} \end{bmatrix}, \tag{2.38}$$

where for each $i = 1, \dots, p$ one has

$$\mathbb{J}_{i1} = \begin{bmatrix} \nabla\omega_i \\ \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{2\times p}, \qquad \mathbb{J}_{i2} = \begin{bmatrix} \boldsymbol{0} \\ (\boldsymbol{e}^{(i)})^T \end{bmatrix} \in \mathbb{R}^{2\times p}. \tag{2.39}$$

Here $\mathbf{0} \in \mathbb{R}^{1\times p}$ and $\boldsymbol{e}^{(i)} \in \mathbb{R}^p$ denotes the ith Euclidean basis vector. Thus, we immediately see that the square matrix $\nabla\Psi$ is surjective iff $\nabla\omega_i$, $i = 1, \ldots, p$, are linearly independent. Unfortunately, this cannot be guaranteed; moreover, considering our particular definition of ω_i (1.26), $\nabla\omega_i = \mathbf{0}$ when $\alpha_{i-1} = \alpha_i = \alpha_{i+1}$. In other words, whenever the contact boundary has a flat part consisting of at least two line segments, $\nabla\Psi$ contains a zero row and, consequently, cannot be surjective.

Computation of D^*Z

In the sequel we will compute the coderivative of Z at a given point $(\bar{x}_1, \bar{x}_2, \bar{z}) \in \operatorname{Gr} Z$. The obtained results will then be used to validate condition (2.36), while at the same time they play a central role in the assertion of Proposition 3 itself.

Let us distinguish several situations according to the position of the reference point $(\bar{x}_1, \bar{x}_2, \bar{z})$ on the graph of Z–see Figure 2.2, where red and green colour mark those points at which *sliding* occurs; the vertical, blue region signifies *sticking*. Points on the common boundary of these sets are said to be in the so-called *weak sticking* mode.

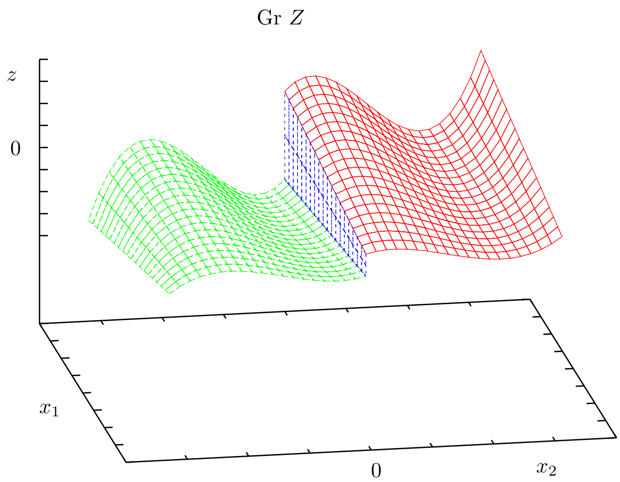


Figure 2.2: Graph of the multifunction $Z(x_1, x_2) = x_1\mathfrak{F}(|x_2|)\partial|x_2|$.

Proposition 4 (sliding)**.** *Let $z^* \in \mathbb{R}$ be arbitrary and $(\bar{x}_1, \bar{x}_2, \bar{z}) \in \operatorname{Gr} Z$ such that $\bar{x}_2 > 0$. Then:*

$$D^*Z(\bar{x}_1, \bar{x}_2, \bar{z})(z^*) = \{z^*\mathfrak{F}(\bar{x}_2)\} \times D^*\mathfrak{F}(\bar{x}_2)(\bar{x}_1 z^*). \tag{2.40}$$

Proof. Due to the assumption on $\bar{x}_2$ there exists a neighbourhood $\mathcal{O}$ of $(\bar{x}_1, \bar{x}_2)$ so that:

$$Z(x_1, x_2) = x_1\mathfrak{F}(x_2) \qquad \forall (x_1, x_2) \in \mathcal{O}.$$

Note that Z is single-valued and (locally) Lipschitz continuous in $\mathcal{O}$. The computation of the regular normal cone to $\operatorname{Gr} Z$ at points of $\mathcal{O}$ is straightforward and yields:

$$\widehat{N}_{\operatorname{Gr} Z}(x_1, x_2, z) = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(x_2), \\ (x_2^*, x_1 z^*) \in \widehat{N}_{\operatorname{Gr}\mathfrak{F}}(x_2, \mathfrak{F}(x_2))\}. \tag{2.41}$$

Thus

$$N_{\mathrm{Gr}\,Z}(\bar{x}_1, \bar{x}_2, \bar{z}) = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(\bar{x}_2), (x_2^*, \bar{x}_1 z^*) \in N_{\mathrm{Gr}\,\mathfrak{F}}(\bar{x}_2, \mathfrak{F}(\bar{x}_2))\},$$

and the assertion follows immediately from the definition of the coderivative. □

Proposition 5 (sliding). *Let $z^* \in \mathbb{R}$ be arbitrary and $(\bar{x}_1, \bar{x}_2, \bar{z}) \in \mathrm{Gr}\, Z$ such that $\bar{x}_2 < 0$. Then:*

$$D^*Z(\bar{x}_1, \bar{x}_2, \bar{z})(z^*) = \{-z^*\mathfrak{F}(-\bar{x}_2)\} \times \big(-D^*\mathfrak{F}(-\bar{x}_2)(-\bar{x}_1 z^*)\big). \tag{2.42}$$

Proof. In this case there exists a neighbourhood $\tilde{\mathcal{O}}$ of $(\bar{x}_1, \bar{x}_2)$ such that:

$$Z(x_1, x_2) = -x_1\mathfrak{F}(-x_2) \qquad \forall (x_1, x_2) \in \tilde{\mathcal{O}}.$$

The rest is done in a similar fashion. □

The previous two cases have the mechanical interpretation of sliding, i.e., represent those contact points, where the displacement in the tangential direction is nonzero.

Proposition 6 (sticking). *Let $z^* \in \mathbb{R}$ be arbitrary and $(\bar{x}_1, 0, \bar{z}) \in \mathrm{Gr}\, Z$ such that $|\bar{z}| < \bar{x}_1\mathfrak{F}(0)$. Then:*

$$D^*Z(\bar{x}_1, 0, \bar{z})(z^*) = \begin{cases} \{0\} \times \mathbb{R} & \text{if } z^* = 0, \\ \emptyset & \text{otherwise.} \end{cases} \tag{2.43}$$

Proof. As readily seen, there exists a neighbourhood $\mathcal{U}$ of $(\bar{x}_1, 0, \bar{z})$ such that:

$$\mathcal{U} \cap \mathrm{Gr}\, Z = \mathcal{U} \cap \big(\mathbb{R} \times \{0\} \times \mathbb{R}\big),$$

whence we immediately get:

$$\widehat{N}_{\mathrm{Gr}\,Z}(x_1, 0, z) = \{0\} \times \mathbb{R} \times \{0\} \qquad \forall (x_1, 0, z) \in \mathcal{U} \cap \mathrm{Gr}\, Z. \tag{2.44}$$

The assertion follows easily from the definition of the coderivative. □

The setting of the previous proposition corresponds to contact points, where (strong) sticking is present, i.e., the tangential component of the stress vector is below the threshold value to trigger motion in the tangential direction. If this critical value is attained at a contact point, but there is still no tangential motion, we speak of weak sticking, which is investigated below.

In order to give a reasonable formula for the coderivative D^*Z at these points, we will, in addition, assume that the coefficient of friction $\mathfrak{F}$ is *weakly semismooth* at 0 (cf. [32]), implying that:

$$\exists \mathfrak{F}'_+(0) \in \mathbb{R} \quad \text{and} \quad \operatorname*{Lim\,sup}_{x \to 0_+} \bar{\partial}\mathfrak{F}(x) = \{\mathfrak{F}'_+(0)\}, \tag{2.45}$$

where $\mathfrak{F}'_+$ stands for the right-hand derivative of $\mathfrak{F}$. Now the following result holds true.

Proposition 7 (weak sticking)**.** *Let $z^* \in \mathbb{R}$ and $\bar{x}_1 > 0$ be arbitrary. Then:*

$$D^*Z(\bar{x}_1, 0, \bar{x}_1\mathfrak{F}(0))(z^*) = \left\{ \begin{bmatrix} z^*\mathfrak{F}(0) \\ \bar{x}_1 z^* \mathfrak{F}'_+(0) + w \end{bmatrix} \middle| \, w \in \begin{cases} \{0\} & \text{if } z^* > 0, \\ \mathbb{R}_- & \text{if } z^* < 0, \\ \mathbb{R} & \text{if } z^* = 0. \end{cases} \right\}. \quad (2.46)$$

Proof. The analysis in this case becomes more involved, since the point $\bar{\boldsymbol{a}} := (\bar{x}_1, 0, \bar{x}_1\mathfrak{F}(0))$ may be approached by sequences corresponding to different mechanical regimes:

$$N_{\operatorname{Gr} Z}(\bar{\boldsymbol{a}}) = \operatorname*{Lim\,sup}_{(x_1,x_2,z) \xrightarrow{\operatorname{Gr} Z} \bar{\boldsymbol{a}}} \widehat{N}_{\operatorname{Gr} Z}(x_1, x_2, z) = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3,$$

where

$$\mathcal{N}_1 := \operatorname*{Lim\,sup}_{\substack{(x_1,x_2,z) \xrightarrow{\operatorname{Gr} Z} \bar{\boldsymbol{a}} \\ x_2 > 0}} \widehat{N}_{\operatorname{Gr} Z}(x_1, x_2, z), \qquad \mathcal{N}_2 := \operatorname*{Lim\,sup}_{\substack{(x_1,0,z) \xrightarrow{\operatorname{Gr} Z} \bar{\boldsymbol{a}} \\ z < x_1\mathfrak{F}(0)}} \widehat{N}_{\operatorname{Gr} Z}(x_1, 0, z),$$

and

$$\mathcal{N}_3 := \operatorname*{Lim\,sup}_{x_1 \to \bar{x}_1} \widehat{N}_{\operatorname{Gr} Z}(x_1, 0, x_1\mathfrak{F}(0)).$$

Observe that the regular normal cones generating in $\mathcal{N}_1$ and $\mathcal{N}_2$ have already been computed in (2.41) and in (2.44), respectively. From (2.44) we immediately have:

$$\mathcal{N}_2 = \{0\} \times \mathbb{R} \times \{0\}.$$

The relation (2.41) may be written as

$$\widehat{N}_{\operatorname{Gr} Z}(x_1, x_2, z) = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(x_2),\ x_2^* \in \widehat{D}^*\mathfrak{F}(x_2)(-\bar{x}_1 z^*)\}. \quad (2.47)$$

Using to the scalarization formula one may write in (2.47):

$$\widehat{D}^*\mathfrak{F}(x_2)(-\bar{x}_1 z^*) \subset D^*\mathfrak{F}(x_2)(-\bar{x}_1 z^*) = \partial(-\bar{x}_1 z^* \mathfrak{F})(x_2) \subset -\bar{x}_1 z^* \bar{\partial}\mathfrak{F}(x_2). \quad (2.48)$$

Taking into account the assumed semismoothness property (3.37), it follows from (2.47) and (2.48)

$$\mathcal{N}_1 = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(x_2),\ x_2^* = -\bar{x}_1 z^* \mathfrak{F}'_+(0)\}, \quad (2.49)$$

since $\mathcal{N}_1 \neq \emptyset$ due to Lipschitz continuity of $\mathfrak{F}$.

The treatment of $\mathcal{N}_3$ is, however, more delicate. As a first step, let us compute the contingent cone to $\operatorname{Gr} Z$ at $\boldsymbol{a} := (x_1, 0, x_1\mathfrak{F}(0))$, for $x_1 > 0$ fixed. Note that $\operatorname{Gr} Z$ locally around the reference point $\boldsymbol{a}$ coincides with the union $G_1 \cup G_2$, where

$$\begin{aligned} G_1 &= \{(x_1', x_2', z') \mid |x_1' - x_1| < \varepsilon,\ x_2' = 0,\ x_1'\mathfrak{F}(0) - \varepsilon < z' \le x_1'\mathfrak{F}(0)\}, \\ G_2 &= \{(x_1', x_2', z') \mid |x_1' - x_1| < \varepsilon,\ 0 \le x_2' < \varepsilon,\ z' = x_1'\mathfrak{F}(x_2')\}, \end{aligned}$$

for a sufficiently small $\varepsilon > 0$. Moreover, the following holds:

$$T_{\operatorname{Gr} Z}(\boldsymbol{a}) = T_{G_1}(\boldsymbol{a}) \cup T_{G_2}(\boldsymbol{a}). \quad (2.50)$$

By the definition of the contingent cone:

$$T_{G_1}(\boldsymbol{a}) = \{(h,k,l) \mid \exists h_i \to h\ \exists k_i \to k\ \exists l_i \to l\ \exists \lambda_i \to 0_+ : \\ \lambda_i k_i = 0,\ x_1\mathfrak{F}(0) + \lambda_i l_i \le (x_1 + \lambda_i h_i)\mathfrak{F}(0)\} = \{(h,0,l) \mid l \le h\mathfrak{F}(0)\}.$$

Analogously:

$$\begin{aligned} T_{G_2}(\boldsymbol{a}) &= \{(h,k,l) \mid \exists h_i \to h\ \exists k_i \to k\ \exists l_i \to l\ \exists \lambda_i \to 0_+ : \\ &\qquad 0 \le \lambda_i k_i,\ x_1\mathfrak{F}(0) + \lambda_i l_i = (x_1 + \lambda_i h_i)\mathfrak{F}(\lambda_i k_i)\} \\ &= \Big\{(h,k,l) \Big| \exists h_i \to h\ \exists k_i \to k\ \exists l_i \to l\ \exists \lambda_i \to 0_+ : \\ &\qquad 0 \le k_i,\ l_i = h_i\mathfrak{F}(\lambda_i k_i) + x_1 k_i \frac{\mathfrak{F}(\lambda_i k_i) - \mathfrak{F}(0)}{\lambda_i k_i}\Big\} \\ &= \{(h,k,l) \mid 0 \le k,\ l = h\mathfrak{F}(0) + x_1 k \mathfrak{F}'_+(0)\}, \end{aligned}$$

where we have made use of (3.37) ensuring directional differentiability of $\mathfrak{F}$ at 0. Now it is sufficient to compute the (negative) polars to these cones to obtain:

$$\widehat{N}_{G_1}(\boldsymbol{a}) = \big(T_{G_1}(\boldsymbol{a})\big)^0 = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(0),\ z^* \ge 0\} \tag{2.51}$$

and similarly:

$$\widehat{N}_{G_2}(\boldsymbol{a}) = \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(0),\ x_2^* \le -x_1 z^* \mathfrak{F}'_+(0)\}. \tag{2.52}$$

Finally, combining (2.50), (2.51) and (2.52) yields:

$$\begin{aligned} \widehat{N}_{\mathrm{Gr}\, Z}(\boldsymbol{a}) &= (T_{G_1}(\boldsymbol{a}) \cup T_{G_2}(\boldsymbol{a}))^0 \\ &= \widehat{N}_{G_1}(\boldsymbol{a}) \cap \widehat{N}_{G_2}(\boldsymbol{a}) \\ &= \{(x_1^*, x_2^*, z^*) \mid x_1^* = -z^*\mathfrak{F}(0),\ x_2^* \le -x_1 z^* \mathfrak{F}'_+(0),\ z^* \ge 0\}. \end{aligned}$$

From this it is obvious that $\mathcal{N}_3 = \widehat{N}_{\mathrm{Gr}\, Z}(\bar{\boldsymbol{a}})$.

In this way we have now an upper estimate of $N_{\mathrm{Gr}\, Z}(\boldsymbol{a})$ and the result follows easily by the definition of the coderivative. Indeed, for instance, the first formula in (2.46) follows from (2.49) and the fact that for $z^* > 0$ and $i = 2, 3$ there does not exist any (x_1^*, x_2^*) such that $(x_1^*, x_2^*, -z^*) \in \mathcal{N}_i$. The statement has been established. □

A straightforward modification of the proof of Proposition 7 implies the following result, concerning the point $\bar{\boldsymbol{a}} := (\bar{x}_1, 0, -\mathfrak{F}(0))$.

Proposition 8 (weak sticking)**.** *Let $z^* \in \mathbb{R}$ and $\bar{x}_1 > 0$ be arbitrary. Then:*

$$D^*Z(\bar{x}_1, 0, -\bar{x}_1\mathfrak{F}(0))(z^*) = \left\{ \begin{bmatrix} -z^*\mathfrak{F}(0) \\ \bar{x}_1 z^* \mathfrak{F}'_+(0) + w \end{bmatrix} \middle| w \in \begin{Bmatrix} \mathbb{R}_+ & \text{if } z^* > 0, \\ \{0\} & \text{if } z^* < 0, \\ \mathbb{R} & \text{if } z^* = 0. \end{Bmatrix} \right\}. \tag{2.53}$$

We are now in a position to verify the qualification condition (2.36).

Corollary 1. *Let $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{q}}) \in \mathrm{Gr}\, \widetilde{Q}_\tau$ be arbitrary. Then (2.36) holds.*

Proof. By (2.40), (2.42), (2.43), (2.46) and (2.53) we see that $D^*Z(\bar{x}_1, \bar{x}_2, \bar{z})(0) \subset \{0\} \times \mathbb{R}$ for any $(\bar{x}_1, \bar{x}_2, \bar{z}) \in \operatorname{Gr} Z$, implying:

$$D^*Z_1(\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}), \bar{\boldsymbol{q}})(0) \subset \big(\{\mathbf{0}\} \times \mathbb{R}\big)^p.$$

Choosing now $\boldsymbol{w} \in \big(\mathbb{R}^2\big)^p$ such that $\boldsymbol{w}_i = (0, c_i)^T$ for all $i = 1, \dots, p$, then (cf. (2.38) and (2.39)):

$$\mathbf{0} = \nabla\Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})^T \boldsymbol{w} = \sum_{i=1}^{p} \nabla\Psi_i(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}})^T \boldsymbol{w}_i = \sum_{i=1}^{p} \begin{bmatrix} \nabla\omega_i(\bar{\boldsymbol{\alpha}})^T & \mathbf{0} \\ \mathbf{0} & \boldsymbol{e}^{(i)} \end{bmatrix} \begin{bmatrix} 0 \\ c_i \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{c} \end{bmatrix}.$$

□

In this way we have proved that the upper estimate (2.37), needed in (AGE), is valid.

The obtained results enable us to establish necessary optimality conditions, that may serve, e.g., as a stopping criterion in the numerical algorithm, or for testing optimality of a design computed in some other way.

Theorem 18. *Let* $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$ *be a local solution to* $(\mathbb{P})$ *(in particular* $\bar{\boldsymbol{y}} = \widetilde{S}(\bar{\boldsymbol{\alpha}})$*). Then:*

(1) $\mathbf{0} \in \nabla_\alpha J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + D^*\widetilde{S}(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})) + N_{\widetilde{U}_{ad}}(\bar{\boldsymbol{\alpha}})$*;*

(2) $\exists\ \boldsymbol{v}^* \in \mathbb{R}^{3p}$*:*

$$\mathbf{0} \in \nabla J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \nabla F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^* + D^*\widetilde{Q}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}, -F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{v}^*) + N_{\widetilde{U}_{ad} \times \mathbb{R}^{3p}}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}).$$

Proof. The optimality condition in (1) amounts directly to the respective condition in [36, Corollary 5.35]. This relation together with Theorem 17 (ii) yields (2). □

Chapter 3

Shape optimization: Coulomb case

In this chapter we consider the optimal shape design problem, where the state problem is given by the two-dimensional Signorini problem with Coulomb friction and a solution-dependent coefficient of friction. Since the analysis of contact problems with Coulomb friction is very involved in the continuous setting, we will investigate existence and computation of discrete optimal shapes only, i.e., for fixed values of the discretization parameter h, and various choices of the cost functional.

The structure of the present chapter copies more or less that of the previous one, with a notable exception: this time only the algebraic setting is considered. First, we shall first derive the reduced version of the algebraic Signorini problem with Coulomb friction and a solution-dependent coefficient of friction and define the shape optimization problem. Then, we prove Lipschitz continuity of the corresponding solution map, but this time using Robinson's strong regularity condition (SRC). As an immediate consequence, one obtains existence of discrete optimal shapes. Moreover, the SRC property will play an important role also in subsequent sensitivity analysis. This is conducted in a similar way as it was done in the previous chapter, using tools from the generalized differential calculus of Mordukhovich. This enables us the efficient solution of the shape optimization problem by means of the ImP and a bundle method of nonsmooth optimization. The results obtained here are have been presented in the paper [5].

3.1 Algebraic shape optimization problem

Let us start with formulating the reduced algebraic state problem. To this end, recall that in Section 1.3.2 we have denoted by $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$ an auxiliary problem representing the algebraic Signorini problem with given friction, where the slip bound is given by the vector $\boldsymbol{g} \in \mathbb{R}^p_+$ and the coefficient of friction by the vector $\mathfrak{F}(\boldsymbol{\varphi}) := [\mathfrak{F}(\varphi_1), \ldots, \mathfrak{F}(\varphi_p)]^T$, for a fixed $\boldsymbol{\varphi} \in \mathbb{R}^p_+$:

$$\left.\begin{aligned}
&\text{Find } (\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p_+ \text{ such that:}\\
&\langle \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}, |\boldsymbol{v}_\tau| - |\boldsymbol{u}_\tau|\rangle_p\\
&\qquad\qquad \geq \langle \boldsymbol{L}(\boldsymbol{\alpha}), \boldsymbol{v} - \boldsymbol{u}\rangle_n + \langle \boldsymbol{\lambda}, \boldsymbol{v}_\nu - \boldsymbol{u}_\nu\rangle_p \quad \forall \boldsymbol{v} \in \mathbb{R}^n,\\
&\langle \boldsymbol{\mu} - \boldsymbol{\lambda}, \boldsymbol{u}_\nu + \boldsymbol{\alpha}\rangle_p \geq 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p_+.
\end{aligned}\right\} \qquad (\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$$

In contrast to the Tresca friction case, there is no reason to restrict the set of admissible design variables, therefore we assume $\boldsymbol{\alpha} \in U_{ad} \subset \mathbb{R}^p_+$ as defined in (1.22), i.e., without imposing additional constraints on the second finite differences.

The pair $(\boldsymbol{u}^*, \boldsymbol{\lambda}^*) \in \mathbb{R}^n \times \mathbb{R}^p_+$ was shown to be the solution of the Signorini problem with Coulomb friction and a solution-dependent coefficient of friction $(M^C(\boldsymbol{\alpha}))$ (in the sense of Definition 9) iff it is the solution to $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}^*, \boldsymbol{g}^*))$, where $(\boldsymbol{\varphi}^*, \boldsymbol{g}^*)$ is a fixed point of $\tilde{\Psi}^C : \mathbb{R}^p_+ \times \mathbb{R}^p_+ \to \mathbb{R}^p_+ \times \mathbb{R}^p_+$, $(\boldsymbol{\varphi}, \boldsymbol{g}) \mapsto (|\boldsymbol{u}_\tau|, \boldsymbol{\lambda})$. On the basis of this relation, we may easily derive the reduced form of $(M^C(\boldsymbol{\alpha}))$ by simply inserting the fixed point of $\tilde{\Psi}^C$ into (1.34). This way one obtains the following system of GEs:

$$\left.\begin{aligned} \mathbf{0} &\in \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\boldsymbol{\alpha}) + \widehat{\boldsymbol{Q}}_\tau(\boldsymbol{u}_\tau, \boldsymbol{\lambda}), \\ \mathbf{0} &= \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\boldsymbol{\alpha}), \\ \mathbf{0} &\in \boldsymbol{u}_\nu + \boldsymbol{\alpha} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}), \end{aligned}\right\} \tag{3.1}$$

where the multifunction $\widehat{\boldsymbol{Q}}_\tau : \mathbb{R}^p \times \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ takes the form

$$\big(\widehat{\boldsymbol{Q}}_\tau(\boldsymbol{v}, \boldsymbol{w})\big)_i = \mathfrak{F}(|v_i|) w_i \partial |v_i| \quad \forall i = 1, 2, \ldots, p.$$

Further, recall that the matrix- and vector-valued mappings $\mathbb{A}_{\tau\tau}, \mathbb{A}_{\tau\nu}, \mathbb{A}_{\nu\tau}, \mathbb{A}_{\nu\nu} : U_{ad} \to \mathbb{R}^{p\times p}$ and $\boldsymbol{L}_\tau, \boldsymbol{L}_\nu : U_{ad} \to \mathbb{R}^p$, respectively, are assumed to be *continuously differentiable.*

Denoting the state variable by $\boldsymbol{y} = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda}) \in \mathbb{R}^{3p}$, we write the system (3.1) in the compact form:

$$\mathbf{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + \widehat{\boldsymbol{Q}}(\boldsymbol{y}), \tag{$GE^C(\boldsymbol{\alpha})$}$$

with $\boldsymbol{F} : U_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}^{3p}$ being the single-valued, continuously differentiable function from the previous chapter (cf. (2.8) and (2.9)):

$$\boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) = \mathcal{A}(\boldsymbol{\alpha})\boldsymbol{y} - \boldsymbol{l}(\boldsymbol{\alpha}).$$

The multivalued mapping $\widehat{\boldsymbol{Q}} : \mathbb{R}^{3p} \rightrightarrows \mathbb{R}^{3p}$ in $(GE^C(\boldsymbol{\alpha}))$ has a closed graph and is given by the expression:

$$\widehat{\boldsymbol{Q}}(\boldsymbol{y}) = \begin{bmatrix} \widehat{\boldsymbol{Q}}_\tau(\boldsymbol{u}_\tau, \boldsymbol{\lambda}) \\ \mathbf{0} \\ N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}) \end{bmatrix} \quad \forall \boldsymbol{y} = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda}) \in \mathbb{R}^{3p}.$$

With the parametrized generalized equation $(GE^C(\boldsymbol{\alpha}))$ we associate the control-to-state mapping $\widehat{\boldsymbol{S}} : U_{ad} \rightrightarrows \mathbb{R}^{3p}$, defined by

$$\widehat{\boldsymbol{S}}(\boldsymbol{\alpha}) := \{\boldsymbol{y} \in \mathbb{R}^{3p} \mid \mathbf{0} \in \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{y}) + \widehat{\boldsymbol{Q}}(\boldsymbol{y})\}.$$

Now the *shape optimization problem* may be stated in the form of the following mathematical program with equilibrium constraints (MPEC):

$$\left.\begin{aligned} \text{minimize} \quad & J(\boldsymbol{\alpha}, \boldsymbol{y}), \\ \text{subj. to} \quad & \boldsymbol{y} \in \widehat{\boldsymbol{S}}(\boldsymbol{\alpha}), \\ & \boldsymbol{\alpha} \in U_{ad}, \end{aligned}\right\} \tag{$\mathbb{P}^C$}$$

where the cost functional $J : U_{ad} \times \mathbb{R}^{3p} \to \mathbb{R}$ is assumed to be *continuously differentiable.* Actually, this smoothness assumption imposed on J is superfluous: as it will become apparent from subsequent analysis, it would be sufficient to assume J Lipschitz continuous. Nevertheless, for ease of presentation, we shall stick to a smooth cost functional, as this does not affect the intrinsic nonsmoothness of $\widehat{\boldsymbol{S}}$, representing the variational inequality $(GE^C(\boldsymbol{\alpha}))$, in any way.

The main result of this section is formulated in the next theorem.

Theorem 19. *Let the assumptions of Theorem 11(ii) hold. Then* $(\mathbb{P}^C)$ *has at least one solution.*

Its proof relies on the compactness of $\mathrm{Gr}\,\widehat{\boldsymbol{S}}$ and will be given below, in a series of auxiliary, but no less important results.

3.2 Lipschitzian stability

Our main aim in this section is to show Lipschitz continuity of $\widehat{\boldsymbol{S}}$. Although one could prove this directly, as it was done in the case of Tresca friction (cf. Theorem 16), the fact that $\widehat{Q}$ does not depend on $\boldsymbol{\alpha} \in U_{ad}$, allows us to prove a stronger result, namely, *strong regularity* of $(GE^C(\boldsymbol{\alpha}))$.

First, however, we shall prove local Lipschitz continuity of the solution to $(M^C(\boldsymbol{\alpha}))$ with respect to the load vector $\boldsymbol{L} \in \mathbb{R}^n$.

3.2.1 Stability with respect to the load vector

Since the domain corresponding to the design vector $\boldsymbol{\alpha} \in U_{ad}$ will be fixed and $\boldsymbol{L}$ variable, let us relabel the problem $(M^C(\boldsymbol{\alpha}))$ by $(M^C(\boldsymbol{L}))$ and the auxiliary problem $(\tilde{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$ by $(\tilde{A}(\boldsymbol{L}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Further, having (1.45) in mind, let:

$$\delta(\|\boldsymbol{L}\|_n) := \frac{\beta + \|\mathbb{A}\|}{\beta\gamma} \max\{C_{max}, R(\|\boldsymbol{L}\|_n) C_{lip}\}. \tag{3.2}$$

We recall from (1.42) the expression for $R(\|\boldsymbol{L}_n\|)$ and define the constant $\kappa > 0$:

$$R(\|\boldsymbol{L}\|_n) = \left[\frac{1}{\gamma} + \frac{1}{\beta}\left(\frac{\|\mathbb{A}\|}{\gamma} + 1\right)\right] \|\boldsymbol{L}\|_n =: \kappa \|\boldsymbol{L}\|_n. \tag{3.3}$$

In terms of the function δ from (3.2), the assumption of Theorem 11(ii) is equivalent to $\delta(\|\boldsymbol{L}\|_n) < 1$. Provided this condition is met, the Signorini problem with Coulomb friction $(M^C(\boldsymbol{\alpha}))$ has a unique solution. In addition, as we will show, the following holds true.

Proposition 9. *Let the assumptions of Theorem 11(ii) be satisfied, i.e.,* $\delta(\|\boldsymbol{L}\|_n) < 1$ *for some* $\boldsymbol{L} \in \mathbb{R}^n$*. Then there exist positive constants* $\epsilon > 0$ *and* $K := K(\boldsymbol{L}, \epsilon) > 0$ *such that:*

$$\|(\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}}) - (\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\lambda}})\|_{n+p} \leq K \|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n \quad \forall \bar{\boldsymbol{L}}, \tilde{\boldsymbol{L}} \in \mathbb{B}_\epsilon(\boldsymbol{L}),$$

where $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}})$, $(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\lambda}})$ *denote the unique solutions of* $(M^C(\bar{\boldsymbol{L}}))$ *and* $(M^C(\tilde{\boldsymbol{L}}))$*, respectively.*

Proof. Existence of $\epsilon > 0$ satisfying:

$$\delta(\|\boldsymbol{L}'\|_n) < 1 \quad \forall \boldsymbol{L}' \in \overline{\mathbb{B}}_\epsilon(\boldsymbol{L}) \tag{3.4}$$

follows immediately by continuity of the function $\delta : \mathbb{R}_+ \to \mathbb{R}_+$ (cf. (3.2)–(3.3)). We choose such an ϵ and denote

$$q := \max\{\delta(\|\boldsymbol{L}'\|_n) \mid \boldsymbol{L}' \in \overline{\mathbb{B}}_\epsilon(\boldsymbol{L})\} \in (0,1).$$

Further, let $\bar{\boldsymbol{L}}, \tilde{\boldsymbol{L}} \in \mathbb{B}_\epsilon(\boldsymbol{L})$ and $(\boldsymbol{\varphi}, \boldsymbol{g}) \in \mathbb{R}^p_+ \times \mathbb{R}^p_+$ be arbitrary. Then, we build the following sequences iteratively:

(i) Let $(\bar{\boldsymbol{u}}^{(0)}, \bar{\boldsymbol{\lambda}}^{(0)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ stand for the solution of the auxiliary problem $(\tilde{A}(\bar{\boldsymbol{L}}, \boldsymbol{\varphi}, \boldsymbol{g}))$. For each $k = 1, 2, 3, \dots$ denote by $(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ the unique solution to $\big(\tilde{A}\big(\bar{\boldsymbol{L}}, |\bar{\boldsymbol{u}}_\tau^{(k-1)}|, \bar{\boldsymbol{\lambda}}^{(k-1)}\big)\big)$.

(ii) Let $(\tilde{\boldsymbol{u}}^{(0)}, \tilde{\boldsymbol{\lambda}}^{(0)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ be the solution of $(\tilde{A}(\tilde{\boldsymbol{L}}, \boldsymbol{\varphi}, \boldsymbol{g}))$. Analogously, for each $k \in \mathbb{N}$ denote by $(\tilde{\boldsymbol{u}}^{(k)}, \tilde{\boldsymbol{\lambda}}^{(k)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ the solution of the problem $\big(\tilde{A}\big(\tilde{\boldsymbol{L}}, |\tilde{\boldsymbol{u}}_\tau^{(k-1)}|, \tilde{\boldsymbol{\lambda}}^{(k-1)}\big)\big)$.

(iii) Finally, for each $k \in \mathbb{N}$ let $(\boldsymbol{U}^{(k)}, \boldsymbol{\Lambda}^{(k)}) \in \mathbb{R}^n \times \mathbb{R}^p_+$ stand for the unique solution of $\big(\tilde{A}\big(\bar{\boldsymbol{L}}, |\tilde{\boldsymbol{u}}_\tau^{(k-1)}|, \tilde{\boldsymbol{\lambda}}^{(k-1)}\big)\big)$.

It follows from the proof of Theorem 11(ii) that the sequences $\big\{\big(|\bar{\boldsymbol{u}}_\tau^{(k)}|, \bar{\boldsymbol{\lambda}}^{(k)}\big)\big\}$, $\big\{\big(|\tilde{\boldsymbol{u}}_\tau^{(k)}|, \tilde{\boldsymbol{\lambda}}^{(k)}\big)\big\}$ tend to the unique fixed-point of $\tilde{\Psi}^C$ in $\mathbb{R}^p_+ \times \mathbb{R}^p_+$, defined in connection with the problems $(M^C(\bar{\boldsymbol{L}}))$ and $(M^C(\tilde{\boldsymbol{L}}))$, respectively. Hence, the sequences $\big\{(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)})\big\}$, $\big\{(\tilde{\boldsymbol{u}}^{(k)}, \tilde{\boldsymbol{\lambda}}^{(k)})\big\}$ converge to the unique solutions $(\bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}})$, $(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\lambda}})$ of $(M^C(\bar{\boldsymbol{L}}))$ and $(M^C(\tilde{\boldsymbol{L}}))$, respectively. Now, making use of Lemma 2(ii), one may write (recall, that we use the norm $\|\boldsymbol{v} + \boldsymbol{w}\|_{r+s} := \|\boldsymbol{v}\|_r + \|\boldsymbol{w}\|_s$ on the product space $\mathbb{R}^r \times \mathbb{R}^s$ for any $r, s \in \mathbb{N}$):

$$\begin{aligned}
&\|(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)}) - (\tilde{\boldsymbol{u}}^{(k)}, \tilde{\boldsymbol{\lambda}}^{(k)})\|_{n+p} \\
&\qquad \le \|(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)}) - (\boldsymbol{U}^{(k)}, \boldsymbol{\Lambda}^{(k)})\|_{n+p} + \|(\boldsymbol{U}^{(k)}, \boldsymbol{\Lambda}^{(k)}) - (\tilde{\boldsymbol{u}}^{(k)}, \tilde{\boldsymbol{\lambda}}^{(k)})\|_{n+p} \\
&\qquad \le \delta(\|\bar{\boldsymbol{L}}\|_n)\big\|\big(|\bar{\boldsymbol{u}}_\tau^{(k-1)}|, \bar{\boldsymbol{\lambda}}^{(k-1)}\big) - \big(|\tilde{\boldsymbol{u}}_\tau^{(k-1)}|, \tilde{\boldsymbol{\lambda}}^{(k-1)}\big)\|_{p+p} + \kappa\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n \\
&\qquad \le q\|(\bar{\boldsymbol{u}}^{(k-1)}, \bar{\boldsymbol{\lambda}}^{(k-1)}) - (\tilde{\boldsymbol{u}}^{(k-1)}, \tilde{\boldsymbol{\lambda}}^{(k-1)})\|_{n+p} + \kappa\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n,
\end{aligned}$$

where κ is from (3.3). Since the above estimate holds for all $k \in \mathbb{N}$, we obtain by induction:

$$\begin{aligned}
&\|(\bar{\boldsymbol{u}}^{(k)}, \bar{\boldsymbol{\lambda}}^{(k)}) - (\tilde{\boldsymbol{u}}^{(k)}, \tilde{\boldsymbol{\lambda}}^{(k)})\|_{n+p} \\
&\quad \le q^k\|(\bar{\boldsymbol{u}}^{(0)}, \bar{\boldsymbol{\lambda}}^{(0)}) - (\tilde{\boldsymbol{u}}^{(0)}, \tilde{\boldsymbol{\lambda}}^{(0)})\|_{n+p} + (q^{k-1} + \dots + q + 1)\kappa\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n \\
&\quad \le q^k\kappa\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n + (q^{k-1} + \dots + q + 1)\kappa\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n \\
&\quad \le \frac{\kappa}{1-q}\|\bar{\boldsymbol{L}} - \tilde{\boldsymbol{L}}\|_n.
\end{aligned}$$

Here we used that $q^k + \dots + q + 1 \le \sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$ for $|q| < 1$, which is satisfied in our case by the definition of q. Now, taking limit as $k \to \infty$ one arrives at the assertion of the proposition. □

Remark 9. Notice, that due to Lemma 6 the solution of $(M(\boldsymbol{\alpha}))$ is *globally* Lipschitz continuous with respect to the load vector, whereas in the Coulomb friction case the same property holds only *locally*.

3.2.2 Strong regularity

Having the above result at hand, we are in a position to prove strong regularity [8, 46] of $(GE^C(\boldsymbol{\alpha}))$.

Proposition 10. *Let the assumption of Theorem 11(ii) hold. Then the generalized equation* $(GE^C(\boldsymbol{\alpha}))$ *is strongly regular at each* $(\boldsymbol{\alpha}, \boldsymbol{y}) \in \operatorname{Gr} \widehat{\boldsymbol{S}}$.

Proof. Let a reference pair $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} \widehat{\boldsymbol{S}}$ be fixed. Recall that $(GE^C(\boldsymbol{\alpha}))$ is called strongly regular at $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})$, provided there exist neighbourhoods $\mathcal{U}$ of $\boldsymbol{0} \in \mathbb{R}^{3p}$ and $\mathcal{V}$ of $\bar{\boldsymbol{y}}$ such, that the mapping:

$$\mathbb{R}^{3p} \ni \boldsymbol{\xi} \mapsto \{\boldsymbol{y} \in \mathcal{V} \mid \boldsymbol{\xi} \in \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \nabla_y \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})(\boldsymbol{y} - \bar{\boldsymbol{y}}) + \widehat{\boldsymbol{Q}}(\boldsymbol{y})\} \tag{3.5}$$

is single-valued and Lipschitz in $\mathcal{U}$. To prove this, let $\boldsymbol{\xi} \in \mathbb{R}^{3p}$ be fixed and notice that $\boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \cdot)$ is linear. Hence, the perturbed GE in (3.5) amounts to: $\boldsymbol{\xi} \in \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \boldsymbol{y}) + \widehat{\boldsymbol{Q}}(\boldsymbol{y})$. The same GE, written componentwise with $\boldsymbol{y} = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu, \boldsymbol{\lambda})$, $\boldsymbol{\xi} = (\boldsymbol{\xi}_\tau, \boldsymbol{\xi}_\nu, \boldsymbol{\xi}_\lambda) \in \mathbb{R}^{3p}$:

$$\left.\begin{aligned}
&\boldsymbol{\xi}_\tau \in \mathbb{A}_{\tau\tau}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\nu - \boldsymbol{L}_\tau(\bar{\boldsymbol{\alpha}}) + \widehat{\boldsymbol{Q}}_\tau(\boldsymbol{u}_\tau, \boldsymbol{\lambda}),\\
&\boldsymbol{\xi}_\nu = \mathbb{A}_{\nu\tau}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\nu - \boldsymbol{\lambda} - \boldsymbol{L}_\nu(\bar{\boldsymbol{\alpha}}),\\
&\boldsymbol{\xi}_\lambda \in \boldsymbol{u}_\nu + \bar{\boldsymbol{\alpha}} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}).
\end{aligned}\right\} \tag{3.6}$$

The system (3.6) may be rewritten as

$$\left.\begin{aligned}
&\boldsymbol{0} \in \mathbb{A}_{\tau\tau}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\tau + \mathbb{A}_{\tau\nu}(\bar{\boldsymbol{\alpha}})(\boldsymbol{u}_\nu - \boldsymbol{\xi}_\lambda) - (\boldsymbol{L}_\tau(\bar{\boldsymbol{\alpha}}) + \boldsymbol{\xi}_\tau - \mathbb{A}_{\tau\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{\xi}_\lambda) + \widehat{\boldsymbol{Q}}_\tau(\boldsymbol{u}_\tau, \boldsymbol{\lambda}),\\
&\boldsymbol{0} = \mathbb{A}_{\nu\tau}(\bar{\boldsymbol{\alpha}})\boldsymbol{u}_\tau + \mathbb{A}_{\nu\nu}(\bar{\boldsymbol{\alpha}})(\boldsymbol{u}_\nu - \boldsymbol{\xi}_\lambda) - \boldsymbol{\lambda} - (\boldsymbol{L}_\nu(\bar{\boldsymbol{\alpha}}) + \boldsymbol{\xi}_\nu - \mathbb{A}_{\nu\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{\xi}_\lambda),\\
&\boldsymbol{0} \in (\boldsymbol{u}_\nu - \boldsymbol{\xi}_\lambda) + \bar{\boldsymbol{\alpha}} + N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}).
\end{aligned}\right\} \tag{3.7}$$

The system of GEs (3.7) represents the Signorini problem with Coulomb friction and a solution-dependent coefficient of friction on the domain given by $\bar{\boldsymbol{\alpha}} \in U_{ad}$ and with load vector

$$\boldsymbol{l}_\xi(\bar{\boldsymbol{\alpha}}) = \begin{bmatrix} \boldsymbol{L}_\tau(\bar{\boldsymbol{\alpha}}) + \boldsymbol{\xi}_\tau - \mathbb{A}_{\tau\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{\xi}_\lambda \\ \boldsymbol{L}_\nu(\bar{\boldsymbol{\alpha}}) + \boldsymbol{\xi}_\nu - \mathbb{A}_{\nu\nu}(\bar{\boldsymbol{\alpha}})\boldsymbol{\xi}_\lambda \\ -\bar{\boldsymbol{\alpha}} \end{bmatrix}, \tag{3.8}$$

having the solution $\boldsymbol{y}_\xi = (\boldsymbol{u}_\tau, \boldsymbol{u}_\nu - \boldsymbol{\xi}_\lambda, \boldsymbol{\lambda})$. As follows from Proposition 9, for suffeciently small $\epsilon > 0$ and $\boldsymbol{\xi} \in \mathcal{U} := \mathbb{B}_\epsilon(\boldsymbol{0})$ the contact problem (3.7) with load vector $\boldsymbol{l}_\xi(\bar{\boldsymbol{\alpha}})$ has exactly one solution, i.e., (3.6) is uniquely solvable. Hence single-valuedness of the mapping (3.5) follows. To see that it is Lipschitz continuous on $\mathcal{U}$, let $\boldsymbol{\xi}^{(1)}, \boldsymbol{\xi}^{(2)} \in \mathcal{U}$ be arbitrary and denote the corresponding solutions of (3.6) by $\boldsymbol{y}^{(1)}$, $\boldsymbol{y}^{(2)}$. Then, employing Proposition 9 ($c > 0$ stands for a generic constant independent of $\boldsymbol{\xi}^{(i)}$, $\boldsymbol{y}^{(i)}$):

$$\begin{aligned}\|\boldsymbol{y}^{(1)}-\boldsymbol{y}^{(2)}\|_{3p} &= \|\boldsymbol{u}_\tau^{(1)}-\boldsymbol{u}_\tau^{(2)}\|_p+\|\boldsymbol{u}_\nu^{(1)}-\boldsymbol{u}_\nu^{(2)}\|_p+\|\boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)}\|_p\\ &\le \|\boldsymbol{u}_\tau^{(1)}-\boldsymbol{u}_\tau^{(2)}\|_p+\|(\boldsymbol{u}_\tau^{(1)}-\boldsymbol{\xi}_\lambda^{(1)})-(\boldsymbol{u}_\tau^{(2)}-\boldsymbol{\xi}_\lambda^{(2)})\|_p\\ &\quad+\|\boldsymbol{\lambda}^{(1)}-\boldsymbol{\lambda}^{(2)}\|_p+\|\boldsymbol{\xi}_\lambda^{(1)}-\boldsymbol{\xi}_\lambda^{(2)}\|_p\\ &\le c\|\boldsymbol{l}_{\xi^{(1)}}(\bar{\boldsymbol{\alpha}})-\boldsymbol{l}_{\xi^{(2)}}(\bar{\boldsymbol{\alpha}})\|_{3p}+\|\boldsymbol{\xi}_\lambda^{(1)}-\boldsymbol{\xi}_\lambda^{(2)}\|_p\\ &\le c\big(\|\boldsymbol{\xi}_\tau^{(1)}-\boldsymbol{\xi}_\tau^{(2)}\|_p+\|\mathbb{A}_{\tau\nu}(\bar{\boldsymbol{\alpha}})\|\|\boldsymbol{\xi}_\lambda^{(1)}-\boldsymbol{\xi}_\lambda^{(2)}\|_p+\|\boldsymbol{\xi}_\nu^{(1)}-\boldsymbol{\xi}_\nu^{(2)}\|_p\\ &\quad+\|\mathbb{A}_{\nu\nu}(\bar{\boldsymbol{\alpha}})\|\|\boldsymbol{\xi}_\lambda^{(1)}-\boldsymbol{\xi}_\lambda^{(2)}\|_p\big)+\|\boldsymbol{\xi}_\lambda^{(1)}-\boldsymbol{\xi}_\lambda^{(2)}\|_p\\ &\le c\|\boldsymbol{\xi}^{(1)}-\boldsymbol{\xi}^{(2)}\|_{3p},\end{aligned}$$

and the proof is complete. □

As a corollary of Proposition 10 we obtain Lipschitz continuity of the corresponding solution map.

Corollary 2. *Let the assumptions of Theorem 11(ii) hold true. Then the solution map $\widehat{\boldsymbol{S}}: U_{ad}\to\mathbb{R}^{3p}$ is single-valued and Lipschitz in U_{ad}.*

Proof. Follows from Theorem 2.1 in [8] and the compactness of U_{ad}. □

Now we are in a position to prove the main result of this section.

Proof of Theorem 19. By Corollary 2 the solution map $\widehat{\boldsymbol{S}}$ is Lipschitz continuous on the compact set U_{ad}, thus its graph is compact in $U_{ad}\times\mathbb{R}^{3p}$. Therefore, any lower semicontinuous cost functional J attains its minimum on $\operatorname{Gr}\widehat{\boldsymbol{S}}$, i.e., the shape optimization problem $(\mathbb{P}^C)$ has at least one solution. □

3.3 Sensitivity analysis

Concerning the numerical solution of the shape optimization problem $(\mathbb{P}^C)$, the same applies as in the Tresca friction case, i.e., due to Corollary 2 we may follow the ImP approach and reformulate the original MPEC into

$$\left.\begin{array}{ll}\text{minimize} & \widehat{\mathcal{J}}(\boldsymbol{\alpha}):=J(\boldsymbol{\alpha},\widehat{\boldsymbol{S}}(\boldsymbol{\alpha}))\\ \text{subj. to} & \boldsymbol{\alpha}\in U_{ad},\end{array}\right\}\qquad(\widehat{\mathbb{P}}^C)$$

where $\widehat{\mathcal{J}}: U_{ad}\to\mathbb{R}$ is (locally) Lipschitz and possibly non-convex. Due to the reasons discussed in Section 2.4 we shall solve $(\widehat{\mathbb{P}}^C)$ by a bundle method. In order to make this approach work, at each step $\bar{\boldsymbol{\alpha}}\in U_{ad}$ of the minimization algorithm one has to be able to provide a function value $\widehat{\mathcal{J}}(\bar{\boldsymbol{\alpha}})=J(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{y}})$ with $\bar{\boldsymbol{y}}=\widehat{\boldsymbol{S}}(\bar{\boldsymbol{\alpha}})$, and one (arbitrary) subgradient $\boldsymbol{\xi}\in\bar{\partial}\widehat{\mathcal{J}}(\bar{\boldsymbol{\alpha}})$. Owing to (2.25)–(2.27) we see that this can be achieved by setting

$$\boldsymbol{\xi}:=\nabla_\alpha J(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{y}})+\boldsymbol{p}^*,$$

where

$$\boldsymbol{p}^*\in D^*\widehat{\boldsymbol{S}}(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}},\bar{\boldsymbol{y}})).$$

The computation of one such $\boldsymbol{p}^*$ is described in the next theorem. Actually, it happens to be a simplified form of Theorem 17(*ii*) for the case when the multifunction $\widehat{Q}$ does not depend on the the design variable $\boldsymbol{\alpha}$.

Theorem 20. *Let* $(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} \widehat{\boldsymbol{S}}$ *be given. Then for each* $\boldsymbol{p}^* \in D^*\widehat{\boldsymbol{S}}(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$ *there exists an adjoint variable* $\boldsymbol{v}^* \in \mathbb{R}^{3p}$ *such that*

$$\boldsymbol{p}^* = \nabla_\alpha F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^* \tag{3.9}$$

and $\boldsymbol{v}^*$ *is a solution to the* adjoint GE*:*

$$\boldsymbol{0} \in \nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) + \nabla_y F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^* + D^*\widehat{Q}(\bar{\boldsymbol{y}}, -F(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{v}^*). \tag{AGE^C}$$

Proof. Due to the strong regularity condition (see Proposition 10) the assumptions of [27, Theorem 5] are satisfied. See also [4, Theorem 4.1]. ☐

Note, that Theorem 20, in general, provides only an upper approximation of $\bar{\partial}\widehat{\mathcal{J}}(\bar{\boldsymbol{\alpha}})$ since the vector $\boldsymbol{v}^*$ constructed using (3.9) and (AGE^C) may lie outside of $D^*\widehat{\boldsymbol{S}}(\bar{\boldsymbol{\alpha}})(\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))$. Let us recall, that this can happen only at points where $\operatorname{Gr} \widehat{Q}$ is *not* graphically regular, and if it does happen (at a nonregular point), the used bundle method may not inevitably collapse. Otherwise a recovery step has to be made in which the bundle method is provided with a *correct* subgradient. Nevertheless, computational experience shows that this occurs very rarely, therefore we will rely on the construction of subgradients via the AGE (AGE^C) as described in Theorem 20.

The rest of this section is devoted to expressing the coderivative $D^*\widehat{Q}$ in terms of the problem data, as $D^*\widehat{Q}$ is the only unknown quantity remaining in (AGE^C). In doing so, we follow closely [4] and begin with reordering the equations in $(GE^C(\boldsymbol{\alpha}))$ so that $\boldsymbol{y} \in (\mathbb{R}^3)^p$ with $\boldsymbol{y}_i = ((\boldsymbol{u}_t)_i, (\boldsymbol{u}_n)_i, \lambda_i)$ comprising all state variables associated with the i-th contact node ($i = 1, \ldots, p$). This way the multifunction $\widehat{Q}$ takes the form:

$$\widehat{Q}(\boldsymbol{y}) = \begin{bmatrix} \Phi(\boldsymbol{y}_1) \\ \Phi(\boldsymbol{y}_2) \\ \vdots \\ \Phi(\boldsymbol{y}_p) \end{bmatrix}, \tag{3.10}$$

where $\Phi : \mathbb{R}^2 \times \mathbb{R}_+ \rightrightarrows \mathbb{R}^3$ is defined as:

$$\Phi(\boldsymbol{a}) := \begin{bmatrix} \mathfrak{F}(|a_1|) a_3 \partial |a_1| \\ 0 \\ N_{\mathbb{R}_+}(a_3) \end{bmatrix} \quad \forall \boldsymbol{a} \in \mathbb{R}^2 \times \mathbb{R}_+. \tag{3.11}$$

Due to the above reordering (3.10) and [47, Example 6.10], one has for every $(\bar{\boldsymbol{y}}, \bar{\boldsymbol{q}}) \in \operatorname{Gr} \widehat{Q}$ and $\boldsymbol{p}^* \in (\mathbb{R}^3)^p$:

$$D^*\widehat{Q}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{q}})(\boldsymbol{p}^*) = \begin{bmatrix} D^*\Phi(\bar{\boldsymbol{y}}_1, \bar{\boldsymbol{q}}_1)(\boldsymbol{p}_1^*) \\ D^*\Phi(\bar{\boldsymbol{y}}_2, \bar{\boldsymbol{q}}_2)(\boldsymbol{p}_2^*) \\ \vdots \\ D^*\Phi(\bar{\boldsymbol{y}}_p, \bar{\boldsymbol{q}}_p)(\boldsymbol{p}_p^*) \end{bmatrix}. \tag{3.12}$$

Therefore, in the sequel we will consider arbitrary $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in \operatorname{Gr} \Phi$, $\boldsymbol{b}^* \in \mathbb{R}^3$ and compute the coderivative $D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*)$ according to the position of $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})$ as given by the following partition of $\operatorname{Gr} \Phi$:

$$\operatorname{Gr} \Phi = L \cup M_1 \cup M_2 \cup M_3^+ \cup M_3^- \cup M_4, \tag{3.13}$$

where the sets on the right-hand side of (3.13) are defined in Table 3.1. From a mechanical point of view, partition (3.13) represents all possible contact and sliding modes of a point on the contact boundary.

<table>
<tr><th></th><th>no contact:
$a_3 = 0,\ b_3 < 0$</th><th>weak contact:
$a_3 = 0,\ b_3 = 0$</th><th>strong contact:
$a_3 > 0,\ b_3 = 0$</th></tr>
<tr><td>sliding:
$a_1 \neq 0$,
$b_1 = \operatorname{sgn}(a_1)\mathfrak{F}(|a_1|)a_3$</td><td rowspan="2">$L$</td><td>$M_2$</td><td>$M_1$</td></tr>
<tr><td>weak sticking:
$a_1 = 0$,
$|b_1| = \mathfrak{F}(0)a_3$</td><td>$M_4$</td><td>$M_3^-$</td></tr>
<tr><td>strong sticking:
$a_1 = 0$,
$|b_1| < \mathfrak{F}(0)a_3$</td><td>$\times\times\times$</td><td>$\times\times\times$</td><td>$M_3^+$</td></tr>
</table>

Table 3.1: Contact and sliding mode at $(\boldsymbol{a}, \boldsymbol{b}) \in \operatorname{Gr}\Phi$.

As easily seen from their definition, the sets L, M_1 and M_3^+ are open in the relative topology of $\operatorname{Gr}\Phi$, i.e., each $\Sigma \in \{L, M_1, M_3^+\}$ satisfies:

$$\forall(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in \Sigma\ \exists\text{neighborhood } \mathcal{O}:\ \operatorname{Gr}\Phi \cap \mathcal{O} \subset \Sigma. \tag{3.14}$$

This makes the analysis in these cases substantially easier, since:

$$N_{\operatorname{Gr}\Phi}(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = N_\Sigma(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = \operatorname*{Lim\,sup}_{(\boldsymbol{a},\boldsymbol{b}) \xrightarrow{\Sigma} (\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})} \widehat{N}_\Sigma(\boldsymbol{a}, \boldsymbol{b}), \tag{3.15}$$

as will be used frequently below.

Proposition 11 (no contact)**.** *Let* $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in L$ *and* $\boldsymbol{b}^* \in \mathbb{R}^3$ *be given. Then:*

$$D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{cases} \{0\} \times \{0\} \times \mathbb{R} & \textit{if } b_3^* = 0, \\ \emptyset & \textit{otherwise.} \end{cases} \tag{3.16}$$

Proof. Let $(\boldsymbol{a}, \boldsymbol{b}) \in L$ be arbitrary. Then there exists a neighborhood $\mathcal{O}$ of $(\boldsymbol{a}, \boldsymbol{b})$ such that:

$$\operatorname{Gr}\Phi \cap \mathcal{O} = \big(\mathbb{R} \times \mathbb{R} \times \{0\}\big) \times \big(\{0\} \times \{0\} \times \mathbb{R}\big) \cap \mathcal{O}.$$

Therefore:

$$\widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a}, \boldsymbol{b}) = \big(\{0\} \times \{0\} \times \mathbb{R}\big) \times \big(\mathbb{R} \times \mathbb{R} \times \{0\}\big), \tag{3.17}$$

and the assertion follows directly from (3.15) and the definition of $D^*\Phi$. ☐

Proposition 12 (strong contact, strong sticking)**.** *Let* $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_3^+$ *and* $\boldsymbol{b}^* \in \mathbb{R}^3$ *be given. Then:*

$$D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{cases} \mathbb{R} \times \{0\} \times \{0\} & \textit{if } b_1^* = 0, \\ \emptyset & \textit{otherwise.} \end{cases} \tag{3.18}$$

Proof. In this case, for every $(\boldsymbol{a},\boldsymbol{b}) \in M_3^+$ one can find a suitable neighborhood $\mathcal{O}$ such that:

$$\operatorname{Gr}\Phi \cap \mathcal{O} = \big(\{0\}\times\mathbb{R}\times\mathbb{R}\big)\times\big(\mathbb{R}\times\{0\}\times\{0\}\big)\cap\mathcal{O},$$

whence

$$\widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a},\boldsymbol{b}) = \big(\mathbb{R}\times\{0\}\times\{0\}\big)\times\big(\{0\}\times\mathbb{R}\times\mathbb{R}\big). \tag{3.19}$$

The rest follows again from (3.15) and the definition of the coderivative. □

Convention. For convenience, in the sequel $\mathfrak{F}$ will signify the *even extension* of the coefficient of friction to the whole $\mathbb{R}$, i.e. $\mathfrak{F}(x) := \mathfrak{F}(-x)\ \forall x < 0$, so that $\mathfrak{F}(|x|) = \mathfrak{F}(x)\ \forall x \in \mathbb{R}$. Clearly, $\mathfrak{F}$ is (globally) Lipschitz in $\mathbb{R}$.

Proposition 13 (strong contact, sliding)**.** *Let* $(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}}) \in M_1$ *and* $\boldsymbol{b}^* \in \mathbb{R}^3$ *be given. Then:*

$$D^*\Phi(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \begin{bmatrix} D^*\mathfrak{F}(\bar{a}_1)(\operatorname{sgn}(\bar{a}_1)\bar{a}_3 b_1^*) \\ 0 \\ \operatorname{sgn}(\bar{a}_1)\mathfrak{F}(\bar{a}_1) b_1^* \end{bmatrix}. \tag{3.20}$$

Proof. There exists a neighborhood $\tilde{\mathcal{O}}$ of $\bar{\boldsymbol{a}}$ such that Φ is single-valued on $\tilde{\mathcal{O}}$ and equals:

$$\Phi(\boldsymbol{a}) = \begin{bmatrix} \operatorname{sgn}(\bar{a}_1)\mathfrak{F}(a_1)a_3 \\ 0 \\ 0 \end{bmatrix} \qquad \forall \boldsymbol{a} \in \tilde{\mathcal{O}}.$$

From the definition of the regular coderivative:

$$\begin{aligned}\widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a},\Phi(\boldsymbol{a})) = \{&(\boldsymbol{a}^*,\boldsymbol{b}^*) \in \mathbb{R}^3\times\mathbb{R}^3 \mid \\ &\langle \boldsymbol{a}^*, \boldsymbol{x}-\boldsymbol{a}\rangle_3 + \langle \boldsymbol{b}^*, \Phi(\boldsymbol{x})-\Phi(\boldsymbol{a})\rangle_3 \le o(\|\boldsymbol{x}-\boldsymbol{a}\|)\ \forall \boldsymbol{x}\},\end{aligned}$$

employing the Lipschitz continuity of $\mathfrak{F}$. A straightforward calculation yields:

$$\begin{aligned}\widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a},\Phi(\boldsymbol{a})) = \{&(\boldsymbol{a}^*,\boldsymbol{b}^*) \mid a_2^* = 0,\ a_3^* = -\operatorname{sgn}(\bar{a}_1)\mathfrak{F}(a_1)b_1^*, \\ &(a_1^*, \operatorname{sgn}(\bar{a}_1)b_1^* a_3) \in \widehat{N}_{\operatorname{Gr}\mathfrak{F}}(a_1,\mathfrak{F}(a_1))\}.\end{aligned} \tag{3.21}$$

Hence (see (3.15)):

$$\begin{aligned}N_{\operatorname{Gr}\Phi}(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}}) = \{&(\boldsymbol{a}^*,\boldsymbol{b}^*) \mid a_2^* = 0,\ a_3^* = -\operatorname{sgn}(\bar{a}_1)\mathfrak{F}(\bar{a}_1)b_1^*, \\ &(a_1^*, \operatorname{sgn}(\bar{a}_1)b_1^*\bar{a}_3) \in N_{\operatorname{Gr}\mathfrak{F}}(\bar{a}_1,\mathfrak{F}(\bar{a}_1))\}\end{aligned}$$

and the proof is complete. □

Remark 10. (i) If $\mathfrak{F}$ happens to be smooth around $\bar{a}_1$, then Φ is smooth in $\tilde{\mathcal{O}}$ and (3.20) reduces to the adjoint Jacobian of Φ, as expected:

$$D^*\Phi(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \operatorname{sgn}(\bar{a}_1)\begin{bmatrix} \mathfrak{F}'(\bar{a}_1)\bar{a}_3 & 0 & 0 \\ 0 & 0 & 0 \\ \mathfrak{F}(\bar{a}_1) & 0 & 0 \end{bmatrix}\boldsymbol{b}^*.$$

(ii) It can be seen from the proofs of Proposition 11 and Proposition 12, that $\operatorname{Gr}\Phi$ is graphically regular at each point of L and M_3^+. It is graphically regular at those points $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_1$ for which $\operatorname{Gr}\mathfrak{F}$ is graphically regular at $(\bar{a}_1, \mathfrak{F}(\bar{a}_1))$. In particular, if $\mathfrak{F}$ is smooth, then $\operatorname{Gr}\Phi$ is graphically regular also on M_1.

Unfortunately, the situation becomes more involved when dealing with the sets M_2 and M_3^-, since they lie on the common boundary of two open sets:

$$M_2 = \operatorname{relint}(\partial L \cap \partial M_1) \quad \text{and} \quad M_3^- = \operatorname{relint}(\partial M_1 \cap \partial M_3^+), \tag{3.22}$$

where $\operatorname{relint}(A)$ denotes the relative interior of the set A.

In order to compute $D^*\Phi$ at points belonging to M_2, we will use a slightly generalized version of [4, Lemma 4.6]. In particular, we show that its assertion holds with equality under less restrictive conditions.

Lemma 9. *Consider a multifunction* $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^o \rightrightarrows \mathbb{R}^l \times \mathbb{R}^p$ *given by*

$$F(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \begin{bmatrix} G(\boldsymbol{x}, \boldsymbol{y}) \\ H(\boldsymbol{y}, \boldsymbol{z}) \end{bmatrix},$$

where $G : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^l$, $H : \mathbb{R}^m \times \mathbb{R}^o \rightrightarrows \mathbb{R}^p$ *are closed-graph multifunctions. Assume that the point* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2)$ *belongs to* $\operatorname{Gr} F$ *and the qualification condition*

$$\left.\begin{aligned} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{w}_2 \end{bmatrix} &\in D^*G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)(\boldsymbol{0}) \\ \begin{bmatrix} -\boldsymbol{w}_2 \\ \boldsymbol{0} \end{bmatrix} &\in D^*H(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)(\boldsymbol{0}) \end{aligned}\right\} \Rightarrow \boldsymbol{w}_2 = \boldsymbol{0} \tag{3.23}$$

holds true. Then one has

$$\begin{aligned} D^*F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_1^*, \boldsymbol{d}_2^*) \subset \{&(\boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{v}_1, \boldsymbol{v}_2) \mid \\ &(\boldsymbol{u}_1, \boldsymbol{u}_2) \in D^*G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)(\boldsymbol{d}_1^*),\ (\boldsymbol{v}_1, \boldsymbol{v}_2) \in D^*H(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_2^*)\}. \end{aligned} \tag{3.24}$$

Assume, in addtion, that for each sequence $\boldsymbol{y}^{(i)} \to \bar{\boldsymbol{y}}$ *and each* $\boldsymbol{\eta} \in D^*G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)(\boldsymbol{d}_1^*)$ *there exist sequences* $(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_1^{(i)}) \xrightarrow{\operatorname{Gr} G} (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)$ *and* $\boldsymbol{d}_1^{*(i)} \to \boldsymbol{d}_1^*$ *such that*

$$\boldsymbol{\eta} \in \operatorname*{Lim\,sup}_{i\to\infty} \widehat{D}^*G(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_1^{(i)})(\boldsymbol{d}_1^{*(i)}). \tag{3.25}$$

Then (3.24) *holds as equality.*

Proof. The first assertion has already been proved in [4]. To prove the second one, let $\boldsymbol{\eta}$ be an element of the right-hand side of (3.24), i.e.

$$\boldsymbol{\eta} = (\boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{v}_1, \boldsymbol{v}_2),$$

for some $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in D^*G(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{f}}_1)(\boldsymbol{d}_1^*)$ and $(\boldsymbol{v}_1, \boldsymbol{v}_2) \in D^*H(\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_2^*)$. Thus, there exist sequences $(\boldsymbol{y}^{(i)}, \boldsymbol{z}^{(i)}, \boldsymbol{f}_2^{(i)}) \xrightarrow{\operatorname{Gr} H} (\bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_2)$, $\boldsymbol{d}_2^{*(i)} \to \boldsymbol{d}_2^*$, $(\boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \to (\boldsymbol{v}_1, \boldsymbol{v}_2)$ such that $(\boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \in \widehat{D}^*H(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_2^{(i)})(\boldsymbol{d}_2^{*(i)})$. By virtue of our additional assumption, there are sequences $\boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}$, $\boldsymbol{f}_1^{(i)} \to \bar{\boldsymbol{f}}_1$, $\boldsymbol{d}_1^{*(i)} \to \boldsymbol{d}_1^*$ and $(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)}) \in \widehat{D}^*G(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{f}_1^{(i)})(\boldsymbol{d}_1^{*(i)})$ such that

$$(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)}) \to (\boldsymbol{u}_1, \boldsymbol{u}_2).$$

It follows from [47, Theorem 10.40] that for all $i \in \mathbb{N}$

$$(\boldsymbol{u}_1^{(i)}, \boldsymbol{u}_2^{(i)} + \boldsymbol{v}_1^{(i)}, \boldsymbol{v}_2^{(i)}) \in \widehat{D}^* F(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}, \boldsymbol{z}^{(i)}, \boldsymbol{f}_1^{(i)}, \boldsymbol{f}_2^{(i)})(\boldsymbol{d}_1^{*(i)}, \boldsymbol{d}_2^{*(i)}),$$

and consequently $\boldsymbol{\eta} \in D^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2)(\boldsymbol{d}_1^*, \boldsymbol{d}_2^*)$. ☐

Remark 11. Note that equality in (3.24) holds also if instead of G the multifunction H satisfies similar conditions as (3.25). The details are left as an easy exercise.

Next we show that the second assumption of Lemma 9, ensuring equality in (3.24), is fulfilled in the case when $\boldsymbol{G}$ is "multiplicatively separable" in the sense that $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y})$, where $f : \mathbb{R}^n \to \mathbb{R}$ is *locally Lipschitz* and $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ is *continuously differentiable*. To this end, let us first present an auxiliary result.

Lemma 10. *Let $n, m, l \in \mathbb{N}$ and the functions $f : \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ be* locally Lipschitz *around $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ and $\bar{\boldsymbol{y}} \in \mathbb{R}^m$, respectively. Let $\boldsymbol{G} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^l$ be defined as*

$$\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) := f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y}).$$

For its regular coderivative then holds:

$$\widehat{D}^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*) = \begin{bmatrix} \widehat{D}^* f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T \boldsymbol{d}^*) \\ \widehat{D}^* \boldsymbol{g}(\bar{\boldsymbol{y}})(f(\bar{\boldsymbol{x}})\boldsymbol{d}^*) \end{bmatrix} \tag{3.26}$$

for any $\boldsymbol{d}^* \in \mathbb{R}^l$.

Proof. From the definition of the regular coderivative we have:

$$\begin{aligned} \widehat{D}^* \boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*) = \{ & (\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathbb{R}^n \times \mathbb{R}^m \mid \\ & \langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n + \langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y}) - f(\bar{\boldsymbol{x}})\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l \\ & \qquad \le o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m) \quad \forall (\boldsymbol{x}, \boldsymbol{y}) \}. \end{aligned}$$

In particular, for $(\boldsymbol{x}, \bar{\boldsymbol{y}})$ and $(\bar{\boldsymbol{x}}, \boldsymbol{y})$ we get the following two relations:

$$\langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n - \langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l \le o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n) \quad \forall \boldsymbol{x}, \tag{3.27}$$

$$\langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})) \rangle_l \le o(\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m) \quad \forall \boldsymbol{y}, \tag{3.28}$$

which immediately yield the inclusion $\subset$ in (3.26) by the definition of the regular coderivative.

To prove the converse inclusion, let us assume that $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{y}^* \in \mathbb{R}^m$ satisfy (3.27) and (3.28), respectively. We sum both equation to get:

$$\begin{aligned} & \langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n + \langle \boldsymbol{y}^*, \boldsymbol{y} - \bar{\boldsymbol{y}} \rangle_m - \langle \boldsymbol{d}^*, f(\boldsymbol{x})\boldsymbol{g}(\boldsymbol{y}) - f(\bar{\boldsymbol{x}})\boldsymbol{g}(\bar{\boldsymbol{y}}) \rangle_l \\ & \qquad \le \langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})) \rangle_l + o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n) + o(\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m). \end{aligned}$$

Finally, to complete the proof, it is sufficient to show that the right-hand side is $o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m)$. The last two terms are left as an easy exercise. Denoting by K_f and K_g the Lipschitz moduli of f and $\boldsymbol{g}$, resp., the first term can be estimated as follows:

$$\frac{\langle \boldsymbol{d}^*, (f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}))(\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}}))\rangle_l}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m}$$
$$\leq \|\boldsymbol{d}^*\|_l \underbrace{\frac{|f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}})|}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n}}_{\leq K_f} \underbrace{\frac{\|\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{g}(\bar{\boldsymbol{y}})\|_l}{\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m}}_{\leq K_g} \underbrace{\frac{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n}{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n + \|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m}}_{\leq 1} \underbrace{\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|_m}_{\to 0} \to 0$$

for $(\boldsymbol{x}, \boldsymbol{y}) \to (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$. □

Remark 12. Notice that the proof of the previous lemma can be applied without change also in case of $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}) := \boldsymbol{f}(\boldsymbol{x})g(\boldsymbol{y})$, where $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^l$ and $g : \mathbb{R}^m \to \mathbb{R}$, both locally Lipschitz. Then one has for every $\boldsymbol{d}^* \in \mathbb{R}^l$:

$$\widehat{D}^*\boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*) = \begin{bmatrix} \widehat{D}^*\boldsymbol{f}(\bar{\boldsymbol{x}})(g(\bar{\boldsymbol{y}})\boldsymbol{d}^*) \\ \widehat{D}^*g(\bar{\boldsymbol{y}})(\boldsymbol{f}(\bar{\boldsymbol{x}})^T\boldsymbol{d}^*) \end{bmatrix}.$$

Proposition 14. *Let the assumptions of Lemma 10 hold, with* $\boldsymbol{g} : \mathbb{R}^m \to \mathbb{R}^l$ continuously differentiable *around* $\bar{\boldsymbol{y}} \in \mathbb{R}^m$. *Then* $\boldsymbol{G}$ *satisfies* (3.25), *i.e.*

$$\forall \boldsymbol{\eta} \in D^*\boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*)\ \forall \boldsymbol{y}^{(i)} \to \bar{\boldsymbol{y}}\ \exists \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}\ \exists \boldsymbol{d}^{(i)} \to \boldsymbol{d}^*\ \exists \boldsymbol{\eta}^{(i)} \to \boldsymbol{\eta} :$$
$$\boldsymbol{\eta}^{(i)} \in \widehat{D}^*\boldsymbol{G}(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})(\boldsymbol{d}^{(i)}).$$

Proof. Let $\boldsymbol{\eta} \in D^*\boldsymbol{G}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{d}^*)$ and $\boldsymbol{y}^{(i)} \to \bar{\boldsymbol{y}}$ be arbitrary. From the scalarization formula and [36, Corollary 1.111(i)] it follows easily that

$$\boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\pi} \\ f(\bar{\boldsymbol{x}})\nabla \boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^* \end{bmatrix} \quad \text{for some} \quad \boldsymbol{\pi} \in D^*f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^*). \tag{3.29}$$

By the definition of the (limiting) coderivative

$$\exists \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}\ \exists r^{(i)} \to \boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^*\ \exists \boldsymbol{\pi}^{(i)} \to \boldsymbol{\pi} : \quad \boldsymbol{\pi}^{(i)} \in \widehat{D}^*f(\boldsymbol{x}^{(i)})(r^{(i)}). \tag{3.30}$$

Let us distinguish between the following two situations.
(i) $\boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^* \neq 0$. Then, clearly, $\boldsymbol{g}(\boldsymbol{y}^{(i)}) \neq \boldsymbol{0}$ for i sufficiently large. For these indices we may select any sequence $\{\boldsymbol{d}^{(i)}\}$ satisfying the conditions

$$\boldsymbol{d}^{(i)} \to \boldsymbol{d}^* \quad \text{and} \quad \boldsymbol{g}(\boldsymbol{y}^{(i)})^T\boldsymbol{d}^{(i)} = r^{(i)}.$$

Observe that such choice of $\{\boldsymbol{d}^{(i)}\}$ is always possible, e.g.

$$\boldsymbol{d}^{(i)} := \frac{r^{(i)}}{\boldsymbol{g}(\boldsymbol{y}^{(i)})^T\boldsymbol{d}^*}\boldsymbol{d}^* \tag{3.31}$$

for i sufficiently large. By Lemma 10

$$\boldsymbol{\eta}^{(i)} := \begin{bmatrix} \boldsymbol{\pi}^{(i)} \\ f(\boldsymbol{x}^{(i)})\nabla \boldsymbol{g}(\boldsymbol{y}^{(i)})^T\boldsymbol{d}^{(i)} \end{bmatrix} \in \widehat{D}^*\boldsymbol{G}(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})(\boldsymbol{d}^{(i)}) \tag{3.32}$$

and so the assertion follows.

(*ii*) $\boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^* = 0$. It follows that $\boldsymbol{\pi} = \boldsymbol{0}$, since $D^*f(\bar{\boldsymbol{x}})(\boldsymbol{g}(\bar{\boldsymbol{y}})^T\boldsymbol{d}^*) = \{\boldsymbol{0}\}$ by virtue of the Mordukhovich criterion [47, Theorem 9.40]. Consider now arbitrary sequences $\boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}$, $\boldsymbol{d}^{(i)} \to \boldsymbol{d}^*$ and $\boldsymbol{\pi}^{(i)} \in \widehat{D}^*f(\boldsymbol{x}^{(i)})(\boldsymbol{g}(\boldsymbol{y}^{(i)})^T\boldsymbol{d}^{(i)}) = D^*f(\boldsymbol{x}^{(i)})(\boldsymbol{g}(\boldsymbol{y}^{(i)})^T\boldsymbol{d}^{(i)}) \neq \emptyset$. Such sequences do exist, because f is differentiable on a dense subset of its domain (Rademacher's theorem) and at these points

$$\widehat{D}^*f(\boldsymbol{x}^{(i)})(r) = D^*f(\boldsymbol{x}^{(i)})(r) \neq \emptyset \quad \forall r \in \mathbb{R}.$$

Clearly, $\boldsymbol{\pi}^{(i)} \to \boldsymbol{0}$ by the outer semicontinuity of the limiting coderivative and the statement follows again from Lemma 10. □

Remark 13. The assertion of Proposition 14 remains valid if we consider $\boldsymbol{G}$ of the form discussed in Remark 12. The only difference is that instead of (3.31) we may take

$$\boldsymbol{d}^{(i)} := \frac{1}{g(\boldsymbol{y}^{(i)})}\boldsymbol{r}^{(i)}.$$

The reader is kindly encouraged to work out the details.

Proposition 15 (weak contact, sliding)**.** *Let* $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_2$ *and* $\boldsymbol{b}^* \in \mathbb{R}^3$ *be given. Then:*

$$D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \left\{ \begin{bmatrix} 0 \\ 0 \\ \operatorname{sgn}(\bar{a}_1)\mathfrak{F}(\bar{a}_1)b_1^* + w \end{bmatrix} \middle| \, w \in \begin{cases} \mathbb{R} & \text{if } b_3^* = 0, \\ \mathbb{R}_- & \text{if } b_3^* < 0, \\ \{0\} & \text{if } b_3^* > 0. \end{cases} \right\}. \tag{3.33}$$

Proof. Consider a reference point $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = (\bar{a}_1, \bar{a}_2, 0, 0, 0, 0) \in M_2$, where $\bar{a}_1 \neq 0$ by the definition of M_2. Then Φ attains the form

$$\Phi(\boldsymbol{a}) = \begin{bmatrix} \operatorname{sgn}(\bar{a}_1)\mathfrak{F}(a_1)a_3 \\ 0 \\ N_{\mathbb{R}_+}(a_3) \end{bmatrix} \qquad \forall \boldsymbol{a} \in \tilde{\mathcal{O}},$$

for a sufficiently small neighborhood $\tilde{\mathcal{O}}$ of $\bar{\boldsymbol{a}}$. Defining the function $G(x, y) := \mathfrak{F}(x)g(y)$, where $g(y) := \operatorname{sgn}(\bar{a}_1)y$ and the closed-graph multifunction $H(y) = N_{\mathbb{R}_+}(y)$, Lemma 9 yields:

$$\begin{aligned} D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \{(u_1, 0, u_2 + v) \mid (u_1, u_2) \in D^*G(\bar{a}_1, 0)(b_1^*), \\ v \in D^*H(0, 0)(b_3^*)\}, \end{aligned} \tag{3.34}$$

because G satisfies the assumptions of Proposition 14 and thus the second assumption of Lemma 9 is satisfied. Since $g(0) = 0$ and $g'(0) = \operatorname{sgn}(\bar{a}_1)$, it follows from (3.29) that

$$D^*G(\bar{a}_1, 0)(b_1^*) = \left\{ \begin{bmatrix} 0 \\ \operatorname{sgn}(\bar{a}_1)\mathfrak{F}(\bar{a}_1)b_1^* \end{bmatrix} \right\}. \tag{3.35}$$

For the coderivative of the normal cone mapping H at $(0, 0) \in \operatorname{Gr} H$ one has:

$$D^*H(0, 0)(b_3^*) = \begin{cases} \mathbb{R} & \text{if } b_3^* = 0, \\ \mathbb{R}_- & \text{if } b_3^* < 0, \\ \{0\} & \text{if } b_3^* > 0. \end{cases} \tag{3.36}$$

Finally, the assertion follows by collecting (3.34), (3.35) and (3.36). □

In order to give a formula for the coderivative $D^*\Phi$ at points in M_3^- we will, in addition, assume that the coefficient of friction $\mathfrak{F}$ is *weakly semismooth* at 0 (cf. [32]), implying that:

$$\exists \mathfrak{F}'_+(0) \in \mathbb{R} \quad \text{and} \quad \operatorname*{Lim\,sup}_{x \to 0_+} \bar{\partial}\mathfrak{F}(x) = \{\mathfrak{F}'_+(0)\}, \tag{3.37}$$

where $\mathfrak{F}'_+$ stands for the right-hand derivative of $\mathfrak{F}$. Now the following result holds true.

Proposition 16 (strong contact, weak sticking)**.** *Let* $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_3^-$ *and* $\boldsymbol{b}^* \in \mathbb{R}^3$ *be given. Then:*

$$D^*\Phi(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})(\boldsymbol{b}^*) = \left\{ \begin{bmatrix} \mathfrak{F}'_+(0)\bar{a}_3 b_1^* + w \\ 0 \\ \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0) b_1^* \end{bmatrix} \,\middle|\, w \in \begin{cases} \mathbb{R} & \text{if } b_1^* = 0, \\ \operatorname{sgn}(\bar{b}_1)\mathbb{R}_+ & \text{if } b_1^* \operatorname{sgn}(\bar{b}_1) < 0, \\ \{0\} & \text{otherwise.} \end{cases} \right\}. \tag{3.38}$$

Proof. Let $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) \in M_3^-$ be given, i.e. $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = (0, \bar{a}_2, \bar{a}_3, \bar{b}_1, 0, 0) \in \mathbb{R}^3 \times \mathbb{R}^3$, where $\bar{a}_3 > 0$ and $|\bar{b}_1| = \mathfrak{F}(0)\bar{a}_3$. It can be easily seen, that there exists a neighborhood $\mathcal{O}$ of $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})$ such that:

$$\operatorname{sgn}(b_1) = \operatorname{sgn}(\bar{b}_1) \quad \text{and} \quad \operatorname{sgn}(a_1)\operatorname{sgn}(\bar{b}_1) \geq 0 \quad \forall (\boldsymbol{a}, \boldsymbol{b}) \in \operatorname{Gr}\Phi \cap \mathcal{O}. \tag{3.39}$$

Moreover (cf. (3.22) and Table 3.1):

$$N_{\operatorname{Gr}\Phi}(\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}}) = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3, \tag{3.40}$$

where

$$\begin{aligned} \mathcal{N}_1 &:= \operatorname*{Lim\,sup}_{(\boldsymbol{a},\boldsymbol{b}) \xrightarrow{M_1} (\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})} \widehat{N}_{M_1}(\boldsymbol{a}, \boldsymbol{b}), \\ \mathcal{N}_2 &:= \operatorname*{Lim\,sup}_{(\boldsymbol{a},\boldsymbol{b}) \xrightarrow{M_3^+} (\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})} \widehat{N}_{M_3^+}(\boldsymbol{a}, \boldsymbol{b}), \\ \mathcal{N}_3 &:= \operatorname*{Lim\,sup}_{(\boldsymbol{a},\boldsymbol{b}) \xrightarrow{M_3^-} (\bar{\boldsymbol{a}}, \bar{\boldsymbol{b}})} \widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a}, \boldsymbol{b}). \end{aligned}$$

Let us first calculate $\mathcal{N}_1$. From (3.21), (3.39) and the definition of the regular coderivative it follows that:

$$\begin{aligned} \widehat{N}_{M_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{x}^*, \boldsymbol{y}^*) \mid x_2^* = 0,\ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0) y_1^*, \\ x_1^* \in \widehat{D}^*\mathfrak{F}(a_1)(-\operatorname{sgn}(\bar{b}_1) a_3 y_1^*)\} \end{aligned} \tag{3.41}$$

for each $(\boldsymbol{a}, \boldsymbol{b}) \in M_1$. Using the scalarization formula and [33, Corollary 3.3.2] we get:

$$\begin{aligned} \widehat{D}^*\mathfrak{F}(a_1)(-\operatorname{sgn}(\bar{b}_1) a_3 y_1^*) &\subset D^*\mathfrak{F}(a_1)(-\operatorname{sgn}(\bar{b}_1) a_3 y_1^*) \\ &= \partial(-\operatorname{sgn}(\bar{b}_1) a_3 y_1^* \mathfrak{F})(a_1) \subset -\operatorname{sgn}(\bar{b}_1) a_3 y_1^* \bar{\partial}\mathfrak{F}(a_1). \end{aligned} \tag{3.42}$$

Note, that $\mathcal{N}_1$ is nonempty (it follows easily from the Lipschitz continuity of $\mathfrak{F}$ and the Rademacher theorem). In light of this fact, (3.41), (3.42) together with the semismoothness assumption (3.37) and (3.39) yield:

$$\begin{aligned}\mathcal{N}_1 = \{(\boldsymbol{a}^*, \boldsymbol{b}^*) \mid a_2^* = 0,\ a_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)b_1^*,\\ a_1^* = -\mathfrak{F}'_+(0)\bar{a}_3 b_1^*\}.\end{aligned} \tag{3.43}$$

Concerning $\mathcal{N}_2$, from (3.19) one has immediately:

$$\mathcal{N}_2 = \big(\mathbb{R} \times \{0\} \times \{0\}\big) \times \big(\{0\} \times \mathbb{R} \times \mathbb{R}\big). \tag{3.44}$$

However, the computation of the cone $\mathcal{N}_3$ is more involved. In particular, let $(\boldsymbol{a}, \boldsymbol{b}) \in M_3^-$ be given and observe that $\operatorname{Gr}\Phi$ locally around $(\boldsymbol{a}, \boldsymbol{b})$ can be written as the union of the following two disjoint sets (cf. Table 3.1 and (3.39)):

$$\begin{aligned}G_1 &:= \{(\boldsymbol{x}, \boldsymbol{y}) \mid \operatorname{sgn}(x_1) = \operatorname{sgn}(\bar{b}_1),\ x_3 > 0,\ y_1 = \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(x_1)x_3,\ y_2 = y_3 = 0\},\\ G_2 &:= \{(\boldsymbol{x}, \boldsymbol{y}) \mid x_1 = 0,\ x_3 > 0,\ \operatorname{sgn}(\bar{b}_1)y_1 \le \mathfrak{F}(0)x_3,\ y_2 = y_3 = 0\}.\end{aligned}$$

This way one has:

$$T_{\operatorname{Gr}\Phi}(\boldsymbol{a}, \boldsymbol{b}) = T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) \cup T_{G_2}(\boldsymbol{a}, \boldsymbol{b}), \tag{3.45}$$

and hence

$$\widehat{N}_{\operatorname{Gr}\Phi}(\boldsymbol{a}, \boldsymbol{b}) = \big(T_{\operatorname{Gr}\Phi}(\boldsymbol{a}, \boldsymbol{b})\big)^0 = \widehat{N}_{G_1}(\boldsymbol{a}, \boldsymbol{b}) \cap \widehat{N}_{G_2}(\boldsymbol{a}, \boldsymbol{b}). \tag{3.46}$$

The contingent cone to G_1 can be determined as follows:

$$\begin{aligned}T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid \exists \boldsymbol{h}^{(i)} \to \boldsymbol{h},\ \boldsymbol{k}^{(i)} \to \boldsymbol{k},\ \lambda^{(i)} \to 0_+,\ \forall i :\\ (\boldsymbol{a} + \lambda^{(i)}\boldsymbol{h}^{(i)}, \boldsymbol{b} + \lambda^{(i)}\boldsymbol{k}^{(i)}) \in G_1\}\\ = \{(\boldsymbol{h}, \boldsymbol{k}) \mid \exists \boldsymbol{h}^{(i)} \to \boldsymbol{h},\ \boldsymbol{k}^{(i)} \to \boldsymbol{k},\ \lambda^{(i)} \to 0_+,\ \forall i :\\ \operatorname{sgn}(\lambda^{(i)}h_1^{(i)}) = \operatorname{sgn}(\bar{b}_1),\ a_3 + \lambda^{(i)}h_3^{(i)} > 0,\\ \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)a_3 + \lambda^{(i)}k_1^{(i)} = \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(\lambda^{(i)}h_1^{(i)})(a_3 + \lambda^{(i)}h_3^{(i)}),\\ \lambda^{(i)}k_2^{(i)} = 0,\ \lambda^{(i)}k_3^{(i)} = 0\},\end{aligned}$$

from which:

$$\begin{aligned}k_1^{(i)} &= \operatorname{sgn}(\bar{b}_1)\frac{\mathfrak{F}(\lambda^{(i)}h_1^{(i)}) - \mathfrak{F}(0)}{\lambda^{(i)}h_1^{(i)}}h_1^{(i)}a_3 + \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(\lambda^{(i)}h_1^{(i)})h_3^{(i)}\\ &= \frac{\mathfrak{F}(\lambda^{(i)}|h_1^{(i)}|) - \mathfrak{F}(0)}{\lambda^{(i)}|h_1^{(i)}|}h_1^{(i)}a_3 + \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(\lambda^{(i)}h_1^{(i)})h_3^{(i)}\\ &\longrightarrow \mathfrak{F}'_+(0)h_1a_3 + \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)h_3, \quad \text{for} \quad i \to \infty,\end{aligned}$$

as follows from (3.37). Thus we get:

$$\begin{aligned}T_{G_1}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid\ \operatorname{sgn}(\bar{b}_1)h_1 \ge 0,\ k_2 = k_3 = 0,\\ k_1 = \mathfrak{F}'_+(0)a_3h_1 + \operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)h_3\}.\end{aligned} \tag{3.47}$$

An analogous computation yields:

$$T_{G_2}(\boldsymbol{a}, \boldsymbol{b}) = \{(\boldsymbol{h}, \boldsymbol{k}) \mid\ h_1 = 0,\ k_2 = k_3 = 0,\ \operatorname{sgn}(\bar{b}_1)k_1 \le \mathfrak{F}(0)h_3\}. \tag{3.48}$$

Now, the negative polars to the cones (3.47), (3.48) can be easily calculated:

$$\begin{aligned}\widehat{N}_{G_1}(\boldsymbol{a},\boldsymbol{b}) = \{(\boldsymbol{x}^*,\boldsymbol{y}^*) \mid\ & (x_1^* + \mathfrak{F}'_+(0)a_3y_1^*)\operatorname{sgn}(\bar{b}_1) \le 0,\\ & x_2^* = 0,\ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)y_1^*\},\end{aligned}$$

and

$$\widehat{N}_{G_2}(\boldsymbol{a},\boldsymbol{b}) = \{(\boldsymbol{x}^*,\boldsymbol{y}^*) \mid x_2^* = 0,\ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)y_1^*,\ y_1^*\operatorname{sgn}(\bar{b}_1) \ge 0\},$$

so that

$$\begin{aligned}\widehat{N}_{G_1}(\boldsymbol{a},\boldsymbol{b}) \cap \widehat{N}_{G_2}(\boldsymbol{a},\boldsymbol{b}) = \{(\boldsymbol{x}^*,\boldsymbol{y}^*) \mid\ & (x_1^* + \mathfrak{F}'_+(0)a_3y_1^*)\operatorname{sgn}(\bar{b}_1) \le 0,\\ & x_2^* = 0,\ x_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)y_1^*,\ y_1^*\operatorname{sgn}(\bar{b}_1) \ge 0\}.\end{aligned} \tag{3.49}$$

Finally, from (3.46) and (3.49) we get:

$$\begin{aligned}\mathcal{N}_3 = \{(\boldsymbol{a}^*,\boldsymbol{b}^*) \mid\ & (a_1^* + \mathfrak{F}'_+(0)\bar{a}_3b_1^*)\operatorname{sgn}(\bar{b}_1) \le 0,\\ & a_2^* = 0,\ a_3^* = -\operatorname{sgn}(\bar{b}_1)\mathfrak{F}(0)b_1^*,\ b_1^*\operatorname{sgn}(\bar{b}_1) \ge 0\}.\end{aligned} \tag{3.50}$$

The assertion of the proposition follows now from (3.43), (3.44), (3.50) and the definition of the coderivative. □

In principle, one could treat the set M_4 (weak contact, weak sticking) in the same way as it was done in Proposition 16 and write the normal cone $N_{\operatorname{Gr}\Phi}(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}})$, $(\bar{\boldsymbol{a}},\bar{\boldsymbol{b}}) \in M_4$, as a union in the sense of (3.40). Some cones in this union are easy to determine, others, however, would require substantially more tedious calculations than it was carried out for $\mathcal{N}_3$ in the previous proof. On the other hand, the set M_4 is merely a 1-dimensional submanifold of the 3-dimensional manifold $\operatorname{Gr}\Phi \subset \mathbb{R}^6$, making it extremely rare to occur in practical computations. From this reason we omit a detailed analysis of M_4 here and do not provide an exact formula for $D^*\Phi$ at these points.

Chapter 4

Numerical realization

In this chapter we will solve the shape optimization problems analyzed in Chapter 2 and Chapter 3. Recall that, following the ImP approach, both shape optimization problems (involving the Tresca and Coulomb laws of friction, respectively) could be formulated as the nonsmooth optimization problem

$$\left.\begin{array}{ll} \text{minimize} & \Theta(\boldsymbol{\alpha}), \\ \text{subj. to} & \boldsymbol{\alpha} \in U, \end{array}\right\} \tag{4.1}$$

where U is a compact subset of $\mathbb{R}^p$ given by linear inequality and/or equality constraints and $\Theta(\boldsymbol{\alpha})$ stands for the composite cost function from $(\widetilde{\mathbb{P}})$ or $(\widehat{\mathbb{P}}^C)$, resp. Either way, Θ is possibly nonconvex and nondifferentiable, due to the intrinsic nonsmoothness of the respective control-to-state mappings $\widetilde{S}$ and $\widehat{S}$, affected partly by the nondifferentiability of the friction coefficient $\mathfrak{F} : [0, \infty) \to (0, \infty)$, as well.

The sensitivity analyses performed in Chapter 2 and Chapter 3, resp., enable us to solve (4.1) with, e.g., a bundle method. From this class of nonsmooth optimization algorithms we have tested the bundle trust [52, 55] and proximal bundle [39] codes. Since both algorithms performed approximately equally well, we chose to introduce the first one in Section 1 of this chapter. At each step $\boldsymbol{\alpha}^{(k)}$, $k \in \mathbb{N}$, the bundle methods need to be supplied with (i) the function value $\Theta(\boldsymbol{\alpha}^{(k)})$ and (ii) one (arbitrary) subgradient from $\bar{\partial}\Theta(\boldsymbol{\alpha}^{(k)})$. The first task involves solving a frictional contact problem with a solution-dependent coefficient of friction—in Section 2 we briefly outline how this can be done. Section 3 is devoted to the second task, in particular, we look at the adjoint equations from Chapters 2 and 3 in more detail. Finally, in Section 4 numerical examples are presented. These were computed by Ing. Petr Beremlijski, Ph.D. using the MatSol [28] library developed at the Technical University in Ostrava.

4.1 The bundle trust method

In this section we briefly outline the main ideas behind the bundle trust (BT) method [52] for the solution of the unconstrained minimization problem

$$\min\{f(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{R}^n\}, \tag{4.2}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is assumed to be *locally Lipschitzian*. Note that additional constraints may be incorporated into (4.2), e.g., via exact penalization.

By the term "bundle methods" one usually refers to a family of related iterative methods for the solution of (4.2) that utilize the *bundle concept* originally introduced by Lemaréchal [29] and Wolfe [54] and have the following features:

- at each iteration $\boldsymbol{x}_k$ a bundle of information $(\boldsymbol{y}_i, f(\boldsymbol{y}_i), \boldsymbol{g}_i) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, $i \in J_k$, is used to build a model of f;
- if the model is not yet adequate, more subgradient information around $\boldsymbol{x}_k$ is integrated into the model.

The first feature is realized by the *cutting plane* approximation of f at $\boldsymbol{x}_k$, i.e., by the piecewise affine function

$$\boldsymbol{x} \mapsto \max_{i \in J_k} \{\boldsymbol{g}_i^T(\boldsymbol{x} - \boldsymbol{y}_i) + f(\boldsymbol{y}_i)\}, \tag{4.3}$$

that equals to f at each $\boldsymbol{y}_i$, $i \in J_k$. Denoting the *linearization error* $\boldsymbol{\alpha}_{k,i} := \boldsymbol{\alpha}(\boldsymbol{x}_k, \boldsymbol{y}_i) = f(\boldsymbol{x}_k) - (\boldsymbol{g}_i^T(\boldsymbol{x}_k - \boldsymbol{y}_i) + f(\boldsymbol{y}_i))$ and introducing the variable $\boldsymbol{d} := \boldsymbol{x} - \boldsymbol{x}_k$ we may express (4.3) as

$$f_{CP}(\boldsymbol{x}_k; \boldsymbol{d}) := \max_{i \in J_k} \{\boldsymbol{g}_i^T \boldsymbol{d} - \boldsymbol{\alpha}_{k,i}\} + f(\boldsymbol{x}_k), \quad \boldsymbol{d} \in \mathbb{R}^n. \tag{4.4}$$

For convex f it holds that $\boldsymbol{\alpha}_{k,i} \geq 0$ for any $k, i \in \mathbb{N}$ and it "measures" the distance of $\boldsymbol{g}_i$ to $\bar{\partial} f(\boldsymbol{x}_k)$ (which amounts in this case to the convex subdifferential), in particular, $\boldsymbol{\alpha}_{k,i} = 0$ iff $\boldsymbol{g}_i \in \bar{\partial} f(\boldsymbol{x}_k)$. This is no longer true for nonconvex f, in which case $\boldsymbol{\alpha}_{k,i}$ is replaced by $\beta_{k,i} := \beta(\boldsymbol{x}_k, \boldsymbol{y}_i) = \max\{\boldsymbol{\alpha}_{k,i}, c_0\|\boldsymbol{x}_k - \boldsymbol{y}_i\|^2\}$, where c_0 is a small positive parameter. This modification ensures that whenever $\boldsymbol{y}_i$ is "far away" from $\boldsymbol{x}_k$, $\beta_{k,i}$ is large and hence $\boldsymbol{g}_i$ plays a minor role in $f_{CP}(\boldsymbol{x}_k; \cdot)$. Again, as the approximation f_{CP} presumably does not model f well far away from $\boldsymbol{x}_k$, one also adds a stabilizing quadratic term $(1/2t_k)\|\boldsymbol{d}\|^2$ to the model, where $t_k > 0$ has still to be chosen appropriately. In BT this is done via a trust region concept while computing the next iterate $\boldsymbol{x}_{k+1}$ from $\boldsymbol{x}_k$. Conceptually, this inner loop may be formulated as follows:

1. compute

$$\boldsymbol{d}_k := \boldsymbol{d}(t_k) = \arg\min \left\{ f_{CP}(\boldsymbol{x}_k; \boldsymbol{d}) + \frac{1}{2t_k}\|\boldsymbol{d}\|^2 \mid \boldsymbol{d} \in \mathbb{R}^n \right\}; \tag{4.5}$$

2. if $f(\boldsymbol{x}_k + \boldsymbol{d}_k)$ is "sufficiently smaller" than $f(\boldsymbol{x}_k)$, then either:

   (a) enlarge t_k and go back to step 1., or

   (b) make a **Serious Step**: set $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \boldsymbol{d}_k$ and compute $\boldsymbol{g}_{k+1} \in \bar{\partial} f(\boldsymbol{x}_{k+1})$;

   if $f(\boldsymbol{x}_k + \boldsymbol{d}_k)$ is "not sufficiently smaller" than $f(\boldsymbol{x}_k)$, then either:

   (a) reduce t_k and go back to step 1., or

   (b) make a **Null Step**: set $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k$ and compute $\boldsymbol{g}_{k+1} \in \bar{\partial} f(\boldsymbol{x}_k + \boldsymbol{d}_k)$.

The quadratic subproblem (4.5) may be equivalently formulated as (ignoring the constant term $f(\boldsymbol{x}_k)$):

$$(v_k, \boldsymbol{d}_k) := \arg\min \left\{ v + \frac{1}{2t_k}\|\boldsymbol{d}\|^2 \mid v \geq \boldsymbol{g}_i^T \boldsymbol{d} - \boldsymbol{\alpha}_{k,i} \ \forall i \in J_k \right\} \in \mathbb{R} \times \mathbb{R}^n. \quad (4.6)$$

Here v_k has the meaning of a predicted decrease in the f based on the approximation f_{CP} around $\boldsymbol{x}_k$. The decision in step 2 of the above algorithm whether to make a serious or null step is then made by comparing v_k with the actual decrease $f(\boldsymbol{x}_k + \boldsymbol{d}_k) - f(\boldsymbol{x}_k)$, provided it is also ensured that the CP-model gets changed substantially when updating the bundle with the computed values. This is made precise in [52], where the complete algorithm may be found.

We conclude this section with the following convergence result (cf. [52]).

Theorem 21. *Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is weakly semismooth, bounded from below and the sequence of iterates $\{\boldsymbol{x}_k\}$ generated by the BT algorithm is bounded. Then $\{\boldsymbol{x}_k\}$ has a C-stationary cluster point $\bar{\boldsymbol{x}}$, i.e., $0 \in \bar{\partial} f(\bar{\boldsymbol{x}})$.*

Semismoothness of our composite cost functional Θ in (4.1) is inherently connected to the semismoothness of the control-to-state mappings $\widetilde{S}$ and $\widehat{S}$, respectively—composition of semismooth functions yields a semismooth function [44]. Unfortunately, the latter property has not been proven so far in either case. At the moment, however, there seem to exist at least two viable ways: (i) prove semismoothness of the control-to-state mappings from the definition as it was done in, e.g., [42], or (ii) by proving and employing a variant of the proposition in [26, Exercise 13] for multifunctions. Nevertheless, a thorough investigation is subject to future research.

4.2 On solving the state problem

Next, we show how the state problems $(M(\boldsymbol{\alpha}))$ and $(M^C(\boldsymbol{\alpha}))$ are solved numerically for a fixed $\boldsymbol{\alpha} \in U_{ad}$. In both cases we utilize the fixed-point approach to reduce our problems to solving a contact problem with given friction and a coefficient that does *not* depend on the solution. Since the overall efficiency depends very much on the fast solution of these subproblems, we briefly describe how it is implemented in MatSol [28].

4.2.1 Outer loop

In both the Tresca and Coulomb friction case we start from their fixed-point formulation, forming the outer loop in the solution algorithm. We shall employ the results and notation from Chapter 1.

Tresca case:

```
choose φ ∈ ℝ^p_+ ,  tol > 0 ,  err > tol
while ( err > tol )
    solve (Ā(α, φ)) to get (u, λ) ∈ ℝ^n × ℝ^p
    update φ := |u_ν|
    update err
end
```

Coulomb case:

```
choose φ, g ∈ ℝ^p_+ ,  tol > 0 ,  err > tol
while ( err > tol )
    solve (Ā(α, φ, g)) to get (u, λ) ∈ ℝ^n × ℝ^p
    update φ := |u_ν| ,  g := λ
    update err
end
```

Note, that both problems $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}))$ and $(\bar{A}(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}))$ represent a Signorini problem with given friction where the coefficient of friction does not depend on the solution as given in $(\bar{A}(\boldsymbol{\alpha}))$ ($\boldsymbol{\alpha}$ is fixed throughout this section). These subproblems are solved iteratively again, as described below.

4.2.2 Inner loop

Instead of solving $(\bar{A}(\boldsymbol{\alpha}))$ in the presented mixed form, the so-called reciprocal approach [18] is used. To this end, one introduces Lagrange multipliers onto the tangential displacement:

$$\Lambda_\tau(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}) := \{\boldsymbol{\mu}_\tau \in \mathbb{R}^p \mid |\boldsymbol{\mu}_\tau| \le \omega(\boldsymbol{\alpha}) \bullet \mathfrak{F}(\boldsymbol{\varphi}) \bullet \boldsymbol{g}\}.$$

Further, let us denote by $\boldsymbol{\lambda}_\nu \in \Lambda_\nu := \mathbb{R}^p_+$ the second component of the solution to $(\bar{A}(\boldsymbol{\alpha}))$, and let $\mathbb{N}, \mathbb{T} \in \mathbb{R}^{p \times n}$ be the matrix representation of the linear mappings $\boldsymbol{u} \mapsto \boldsymbol{u}_\nu$ and $\boldsymbol{u} \mapsto \boldsymbol{u}_\tau$, respectively. This way $(\bar{A}(\boldsymbol{\alpha}))$ may be equivalently written as:

$$\left.\begin{array}{l} \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u} + \mathbb{T}^T\boldsymbol{\lambda}_\tau = \boldsymbol{L}(\boldsymbol{\alpha}) + \mathbb{N}^T\boldsymbol{\lambda}_\nu, \\ \langle \boldsymbol{\mu}_\tau - \boldsymbol{\lambda}_\tau, \mathbb{T}\boldsymbol{u}\rangle_p + \langle \boldsymbol{\mu}_\nu - \boldsymbol{\lambda}_\nu, \mathbb{N}\boldsymbol{u} + \boldsymbol{\alpha}\rangle_p \le 0 \quad \forall(\boldsymbol{\mu}_\tau, \boldsymbol{\mu}_\nu) \in \Lambda_\tau(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}) \times \Lambda_\nu. \end{array}\right\} \tag{4.7}$$

One arrives at the dual formulation of $(\bar{A}(\boldsymbol{\alpha}))$ after eliminating the primal variable $\boldsymbol{u} \in \mathbb{R}^n$ from the system above. The resulting variational inequality is equivalent to

$$\left.\begin{array}{rl} \text{minimize} & \frac{1}{2}\langle \mathbb{Q}(\boldsymbol{\alpha})\boldsymbol{\mu}, \boldsymbol{\mu}\rangle_{2p} - \langle \boldsymbol{H}(\boldsymbol{\alpha}), \boldsymbol{\mu}\rangle_{2p}, \\ \text{subj. to} & \boldsymbol{\mu} = (\boldsymbol{\mu}_\tau, \boldsymbol{\mu}_\nu) \in \Lambda_\tau(\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{g}) \times \Lambda_\nu, \end{array}\right\} \tag{4.8}$$

where

$$\mathbb{Q}(\boldsymbol{\alpha}) := \begin{bmatrix} \mathbb{T}\mathbb{A}^{-1}(\boldsymbol{\alpha})\mathbb{T}^T & -\mathbb{T}\mathbb{A}^{-1}(\boldsymbol{\alpha})\mathbb{N}^T \\ -\mathbb{N}\mathbb{A}^{-1}(\boldsymbol{\alpha})\mathbb{T}^T & \mathbb{N}\mathbb{A}^{-1}(\boldsymbol{\alpha})\mathbb{N}^T \end{bmatrix}, \quad \boldsymbol{H}(\boldsymbol{\alpha}) := \begin{bmatrix} \mathbb{T}\mathbb{A}^{-1}(\boldsymbol{\alpha})\boldsymbol{L}(\boldsymbol{\alpha}) \\ -\mathbb{N}\mathbb{A}^{-1}(\boldsymbol{\alpha})\boldsymbol{L}(\boldsymbol{\alpha}) \end{bmatrix}.$$

Not only is the dimension of (4.8) considerably less than in case of $(\bar{A}(\boldsymbol{\alpha}))$ ($p \ll n$; the dual variables relate to the contact boundary only), but there exist efficient methods for its solution. The MatSol library implements a conjugate gradient method with proportioning and projections [9] (see also [10]) for solving the quadratic problem (4.8) with simple (box-) constraints.

4.3 On solving the adjoint generalized equation

In this section we shall revisit the adjoint generalized equations (AGE) and (AGE^C) which are supposed to yield a subgradient of the cost functional. Based on the results of Chapter 2 and 3 we will make their solution more obvious.

4.3.1 Tresca case

In Section 2.5 we have argued that a subgradient of the composite cost functional $\mathcal{J}$ can be conveniently approximated by solving (AGE) for $\boldsymbol{p}^* \in \mathbb{R}^p$ and inserting it into (2.27). The idea behind solving (AGE) is to identify a linear subspace in D^*Q for which the resulting system of linear equations can be easily solved. In order to do so we combine the results obtained in Section 2.5, proceeding in reverse order.

Let $\bar{\boldsymbol{\alpha}} \in \widetilde{U}_{ad}$ and the corresponding state vector $\bar{\boldsymbol{y}} = (\bar{\boldsymbol{u}}_\nu, \bar{\boldsymbol{u}}_\tau, \bar{\boldsymbol{\lambda}}) := \widetilde{S}(\bar{\boldsymbol{\alpha}})$ be given. Based on the type of sliding/sticking at the ith contact node and relations (2.40), (2.42), (2.43), (2.46), (2.53), we determine at each contact node $i = 1, \ldots, p$ a linear subspace

$$\mathcal{L}_i \subset D^*Z\big(\omega_i(\bar{\boldsymbol{\alpha}}), (\bar{\boldsymbol{u}}_\tau)_i, -(\boldsymbol{F}_1(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))_i\big), \tag{4.9}$$

i.e., we either choose $v_i^* = 0$ or there exist $d_i^{(1)}, d_i^{(2)} \in \mathbb{R}$ such that for

$$\mathcal{L}_i := \{(a_i^*, b_i^*, v_i^*) \in \mathbb{R}^3 \mid a_i^* = d_i^{(1)} v_i^*,\ b_i^* = d_i^{(2)} v_i^*\} \tag{4.10}$$

(4.9) holds. In the former case we simply omit the equation corresponding to the index i from (AGE), therefore let us assume that the latter case holds for each $i = 1, \ldots, p$. For later use we denote the vectors $\boldsymbol{a}^* := (a_1^*, \ldots, a_p^*)^T, \boldsymbol{b}^* := (b_1^*, \ldots, b_p^*)^T \in \mathbb{R}^p, \boldsymbol{z}^* := ((a_1^*, b_1^*), \ldots, (a_p^*, b_p^*))^T \in (\mathbb{R}^2)^p$ and the diagonal matrices $\mathbb{D}^{(1)}, \mathbb{D}^{(2)} \in \mathbb{R}^{p \times p}$ having the values $d_i^{(1)}$ and $d_i^{(2)}$ as their diagonal entries, respectively, so that

$$\boldsymbol{a}^* = \mathbb{D}^{(1)} \boldsymbol{v}^* \quad \text{and} \quad \boldsymbol{b}^* = \mathbb{D}^{(2)} \boldsymbol{v}^*. \tag{4.11}$$

From (2.37), (2.38), (2.39) and (4.11) we infer that

$$\boldsymbol{\zeta} := \nabla \Psi(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}_\tau)^T \boldsymbol{z}^* = \begin{bmatrix} \nabla \omega(\bar{\boldsymbol{\alpha}})^T \boldsymbol{a}^* \\ \boldsymbol{b}^* \end{bmatrix} = \begin{bmatrix} \nabla \omega(\bar{\boldsymbol{\alpha}})^T \mathbb{D}^{(1)} \\ \mathbb{D}^{(2)} \end{bmatrix} \boldsymbol{v}_1^* \tag{4.12}$$

approximates a vector in $D^*\widetilde{Q}_\tau(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{u}}_\tau, -F_1(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{v}^*)$. This yields the first component of the coderivative D^*Q in (2.34).

Similarly, a vector $\boldsymbol{c}^* \in D^* N_{\mathbb{R}_+^p}(\bar{\boldsymbol{\lambda}}, -\boldsymbol{F}_3(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}))(\boldsymbol{w}^*)$, $\boldsymbol{w}^* \in \mathbb{R}^p$ arbitrary, can be constructed on the basis of Proposition 2 as follows. At each contact point $i \in \{1, \ldots, p\}$ we determine the type of contact:

- if there is no contact (cf. Prop. 2(i)), we set $c_i^* = 0$;
- if there is strong contact (cf. Prop. 2(ii)) we set $w_i^* = 0$ and exclude the corresponding equation from (AGE);
- for weak contact (cf. Prop. 2(iii)) we decide for one of the options described above.

For simplicity of notation, let us assume that $c_i^* = 0$ holds for each $i = 1, \ldots, p$, i.e., $\boldsymbol{c}^* = \boldsymbol{0} \in \mathbb{R}^p$.

Now, writing the adjoint Jacobian of $\boldsymbol{F}$ as

$$\nabla \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T = \begin{bmatrix} \nabla_\alpha \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \\ \nabla_y \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \end{bmatrix} = \begin{bmatrix} \nabla_\alpha (\mathcal{A}(\bar{\boldsymbol{\alpha}})\bar{\boldsymbol{y}})^T - \nabla l(\bar{\boldsymbol{\alpha}})^T \\ \mathcal{A}^T(\bar{\boldsymbol{\alpha}}) \end{bmatrix} \tag{4.13}$$

we compute a solution of (AGE) by solving the system of linear equations:

$$\boldsymbol{p}^* = \left(\nabla_\alpha (\mathcal{A}(\bar{\boldsymbol{\alpha}})\bar{\boldsymbol{y}}) - \nabla l(\bar{\boldsymbol{\alpha}}) + \mathcal{D}^{(1)} \nabla \omega(\bar{\boldsymbol{\alpha}})\right)^T \boldsymbol{v}^*, \tag{4.14}$$

$$-\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) = \left(\mathcal{A}(\bar{\boldsymbol{\alpha}}) + \mathcal{D}^{(2)}\right)^T \boldsymbol{v}^*, \tag{4.15}$$

where

$$\mathcal{D}^{(1)} = \begin{bmatrix} \mathbb{D}^{(1)} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{3p \times p} \quad \text{and} \quad \mathcal{D}^{(2)} = \begin{bmatrix} \mathbb{D}^{(2)} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{3p \times 3p}.$$

First, (4.15) is solved for $\boldsymbol{v}^* \in \mathbb{R}^{3p}$, which is then inserted into (4.14) to get the desired vector $\boldsymbol{p}^* \in \mathbb{R}^p$.

Finally, let us comment on the solvability of (4.15). By assumption, the matrix $\mathcal{A}(\bar{\boldsymbol{\alpha}})$ is positive definite for each $\bar{\boldsymbol{\alpha}} \in \widetilde{U}_{ad}$ and the elements of $\mathcal{D}^{(2)}$ are bounded by C_{max} and C_{lip}, which can be made arbitrarily small. Thus, (4.15) is solvable provided $\mathfrak{F}$ is bounded and Lipschitzian with sufficiently small constants C_{max} and C_{lip}.

4.3.2 Coulomb case

Since the solution of (AGE^C) is done in exactly the same way as described in the previous section, let us only highlight the common and distinct features of solving (AGE^C) in Theorem 20.

Comparing with (4.14), we immediately see that (3.9) does not contain an additional term coming from the coderivative of multifunction $\widehat{Q}$. This follows from the fact that $\widehat{Q}$ does not depend on the design variable $\boldsymbol{\alpha}$. The GE (AGE^C) is treated analogously to the Tresca case: based on (3.12) and the expressions in Propositions 11—16 one assembles the matrix

$$\widehat{\mathcal{D}} = \begin{bmatrix} \widehat{\mathbb{D}}^{(1)} & \boldsymbol{0} & \widehat{\mathbb{D}}^{(2)} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{3p \times 3p},$$

where the entries of the diagonal matrices $\widehat{\mathbb{D}}^{(j)} \in \mathbb{R}^{p \times p}$, $j = 1, 2$, are again bounded by C_{max} and C_{lip}. Note that in the Coulomb friction law the tangential stress depends also on the normal stress—the third component of our state vector—explaining the presence of $\widehat{\mathbb{D}}^{(2)}$ in the matrix $\widehat{\mathcal{D}}$ (compare with $\mathcal{D}^{(2)}$ from the previous section).

Since the single-valued part of $(GE^C(\boldsymbol{\alpha}))$ coincides with that of (2.8), we can use (4.13) to transform the adjoint system in Theorem 20 into the computationally managable form

$$\boldsymbol{p}^* = \nabla_\alpha \boldsymbol{F}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}})^T \boldsymbol{v}^*, \tag{4.16}$$

$$-\nabla_y J(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{y}}) = \left(\mathcal{A}(\bar{\boldsymbol{\alpha}}) + \widehat{\mathcal{D}}\right)^T \boldsymbol{v}^*. \tag{4.17}$$

Concerning the solvability of (4.17) the same applies as for (4.15).

4.4 Examples

In computations we use a slightly different definition of the discrete admissible set $\mathcal{U}_{ad}^h$. The reason for this is twofold:

(i) to reduce the dimension of the control variables, and

(ii) to obtain a smooth contact boundary $\Gamma_C(\alpha_h)$.

To this end we define $\mathcal{U}_{ad}^h$ as a suitable subset of Bézier functions of order d. Let us recall that Bézier functions (of order d) are defined as

$$B_{\boldsymbol{\alpha}}(x) := \sum_{i=0}^{d} \alpha_i \beta_{d,i}(x), \quad \text{where} \quad \beta_{d,i}(x) := \frac{1}{a^d}\binom{d}{i} x^i (a-x)^{d-i}, \quad x \in [0,a]$$

and $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_d) \in \mathbb{R}^{d+1}$. The points $(ih', \alpha_i) \in \mathbb{R}^2$, $i = 0, \ldots, d$ $(h' := a/d)$ are called the *control points* of $B_{\boldsymbol{\alpha}}$. It holds that $B_{\boldsymbol{\alpha}}(0) = \alpha_0$, $B_{\boldsymbol{\alpha}}(a) = \alpha_d$ and $\operatorname{Gr} B_{\boldsymbol{\alpha}}$ lies in the convex hull of its control points. Moreover, taking the control variable $\boldsymbol{\alpha}$ from the set

$$\begin{aligned} U := \{\boldsymbol{\alpha} \in \mathbb{R}^{d+1} \mid 0 \le \alpha_i \le C_0 \quad \forall i = 0, \ldots, d, \\ |\alpha_{i-1} - \alpha_i| \le C_1 h' \quad \forall i = 1, \ldots, d, \\ |\alpha_{i-1} - 2\alpha_i + \alpha_{i+1}| \le C_3 (h')^2 \quad \forall i = 1, \ldots, d-1, \\ C_{21} \le \int_0^a B_{\boldsymbol{\alpha}}(x)\,dx \le C_{22}\} \end{aligned} \tag{4.18}$$

ensures that the corresponding Bézier function $B_{\boldsymbol{\alpha}}$ satisfies all constraints introduced in (2.5), in particular, $|B'_{\boldsymbol{\alpha}}| \le C_1$ and $|B''_{\boldsymbol{\alpha}}| \le C_3$ everywhere in $[0,a]$. The domain $\Omega(\boldsymbol{\alpha})$ is first approximated by a polygonal one, then triangulated using quadrilaterals to obtain the *computational domain* $\Omega_h(\boldsymbol{\alpha})$. The discrete function spaces on $\Omega_h(\boldsymbol{\alpha})$ are defined using Q_1-isoparametric finite elements of Lagrange type. In all three examples presented below the values (cf. Figure 1.2) $a = 2$, $b = 1$, $d = 20$ are used and the total number of nodes (vertices of quadrilaterals) equals 1800 for each $\boldsymbol{\alpha} \in U$, including 60 on the contact part.

Example 1

First, we will assume the model with Tresca friction and a solution-dependent coefficient of friction $\mathfrak{F}$, given by the smooth function

$$\mathfrak{F}(s) := 0.25 \frac{1}{1+s^2}, \quad s \in \mathbb{R}_+. \tag{4.19}$$

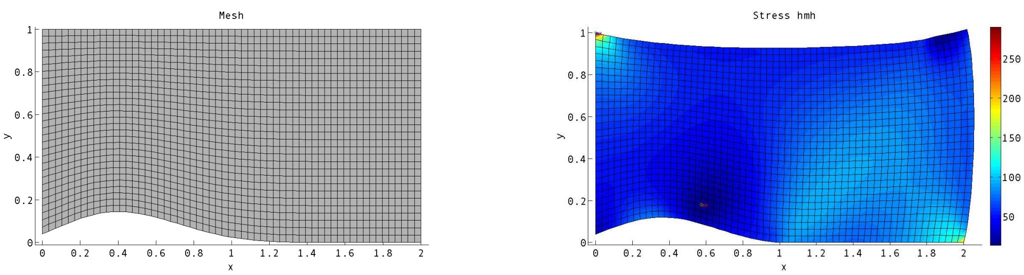
The slip bound shall take the constant value $g = 150$. In the present example we will try to identify the contact normal stress $\boldsymbol{\lambda}$ by a prescribed target value $\boldsymbol{\lambda}_{\text{tar}}$, as denoted by the dotted line in Figure 4.3. Thus, the discretized shape optimization problem reads as

$$\left.\begin{array}{rl} \text{minimize} & \|\boldsymbol{\lambda}(\boldsymbol{\alpha}) - \boldsymbol{\lambda}_{\text{tar}}\|^2, \\ \text{subj. to} & \boldsymbol{\alpha} \in U, \end{array}\right\}$$

where $\boldsymbol{\lambda}(\boldsymbol{\alpha})$ is the second component of the solution to $(M(\boldsymbol{\alpha}))$. The other parameters in the model were set to the following values: $C_0 = 0.75$, $C_1 = 0.85$, $C_3 = 10$, $C_{21} = 1.88$, $C_{22} = 1.95$; we take a material with Young's modulus $E = 1$ GPa and Poisson constant $\sigma = 0.3$; density of forces that press on the upper edge is $\boldsymbol{P}^1 = (0, -60 \text{ MPa})$ on $(0, 1.8) \times \{1\}$ and zero on $(1.8, 2) \times \{1\}$, while a pulling force of density $\boldsymbol{P}^2 = (50 \text{ MPa}, 30 \text{ MPa})$ acts on the right edge; the body is clamped along its left edge.

The initial design is presented in Figure 4.1 in its unloaded state (left) and the distribution of the von Mises stress in the deformed body (right). Similarly, Figure 4.2 shows the optimal design before and after loading. On Figure 4.3 we compare the normal contact stresses with the prescribed function: while the initial contact stress is far from the target values, the stresses for the optimal design follow $\boldsymbol{\lambda}_{\text{tar}}$ very closely. Let us mention, that the BT algorithm converged from $\boldsymbol{\alpha}_0$ to $\boldsymbol{\alpha}_{\text{opt}}$ in about 150 iterations and the initial value $\mathcal{J}(\boldsymbol{\alpha}_0) = 5.9 \cdot 10^4$ of the cost functional dropped by two orders of magnitude to $\mathcal{J}(\boldsymbol{\alpha}_{\text{opt}}) = 9.1 \cdot 10^2$.

In order to emphasize the importance of proper modelling of contact problems,



(a) Before deformation

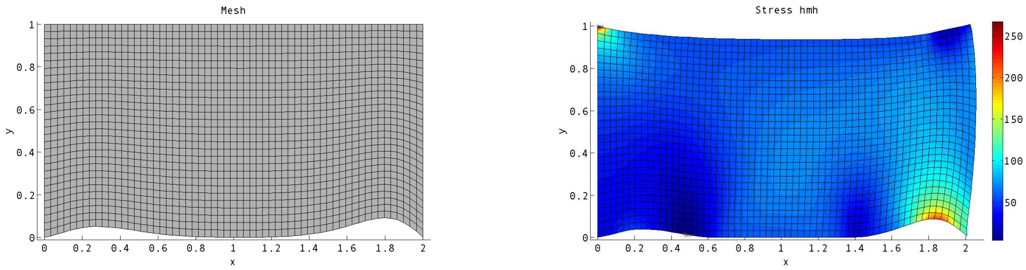
(b) After deformation

Figure 4.1: Example 1; initial design.

let us re-compute the previous example with the following modification: instead of allowing $\mathfrak{F}$ to depend on the unknown solution we fix its value to

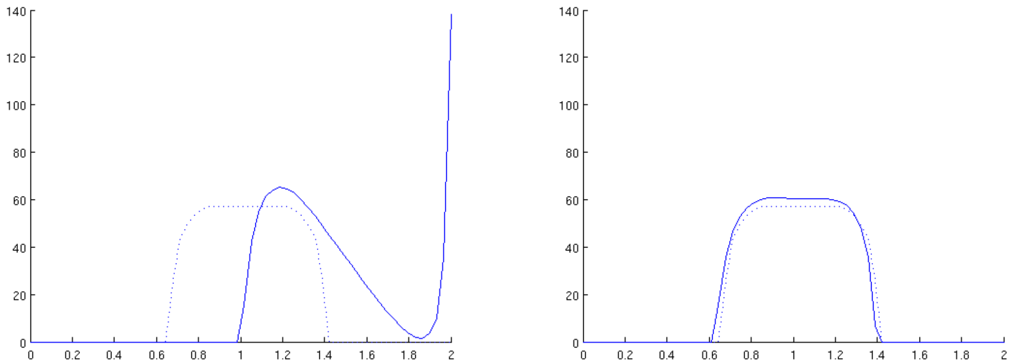
$$\mathfrak{F}(s) := 0.25, \quad s \in \mathbb{R}_+,$$

but keep all other parameters of Example 1 unchanged. Starting from the same initial domain $\Omega(\boldsymbol{\alpha}_0)$, the BT algorithm converges to a solution $\Omega(\bar{\boldsymbol{\alpha}}_{\text{opt}})$–cf. Figure 4.4. At first sight, Figure 4.2 yields a satisfactory correspondence with the



(a) Before deformation (b) After deformation

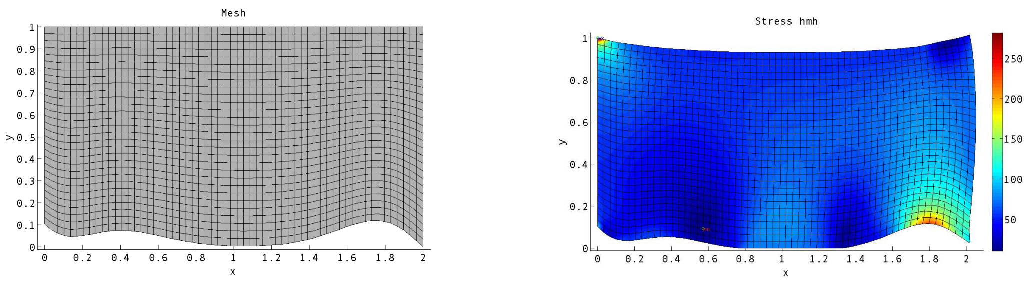
Figure 4.2: Example 1; optimal design.



(a) Initial design (b) Optimal design

Figure 4.3: Example 1; normal stresses.

solution of the unsimplified problem. However, recomputing the original contact problem with (4.19) on $\Omega(\bar{\boldsymbol{\alpha}}_{\text{opt}})$ reveals that $\Omega(\bar{\boldsymbol{\alpha}}_{\text{opt}})$ is actually far from being optimal (cf. Figure 4.5).



(a) Before deformation (b) After deformation

Figure 4.4: Example 1 with $\mathfrak{F} = \text{const}$; optimal design $\Omega(\bar{\boldsymbol{\alpha}}_{\text{opt}})$.

Example 2

In the next two example computations we will consider the contact problems with Coulomb friction ($M^C(\boldsymbol{\alpha})$), but with a much more complicated friction coefficient

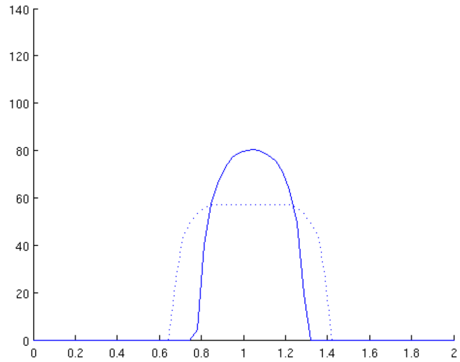


Figure 4.5: Example 1; normal stress distribution on $\Gamma_C(\bar{\boldsymbol{\alpha}}_{\mathrm{opt}})$.

$\mathfrak{F}$, namely

$$\mathfrak{F}(s) := \begin{cases} 0.2 + s, & \text{if } s \in [0, 0.05], \\ 0.25 - s, & \text{if } s \in (0.05, 0.2], \\ 0.1, & \text{if } s \in (0.2, \infty). \end{cases}$$

Note that the function $\mathfrak{F}$ is Lipschitz with modulus 1, but non-differentiable at 0.05 and 0.2, and also non-monotone.

Our aim is to find a suitable contact part, among the ones specified by U, which minimizes peaks of the (discrete) normal contact stress $\lambda_h(\boldsymbol{\alpha})$ represented by the vector of Lagrange multipliers $\boldsymbol{\lambda}(\boldsymbol{\alpha})$, $\boldsymbol{\alpha} \in U$. Since the max-norm $\|\boldsymbol{\lambda}\|_\infty = \max_{i=1,\dots,p} |\lambda_i|$ is not continuously differentiable, we shall use the l^q-norm $|\boldsymbol{\lambda}|_q = (\sum_{i=1}^p |\lambda_i|^q)^{1/q}$ instead, with q large enough ($q = 6$ in our case). Thus, the shape optimization problem reads as

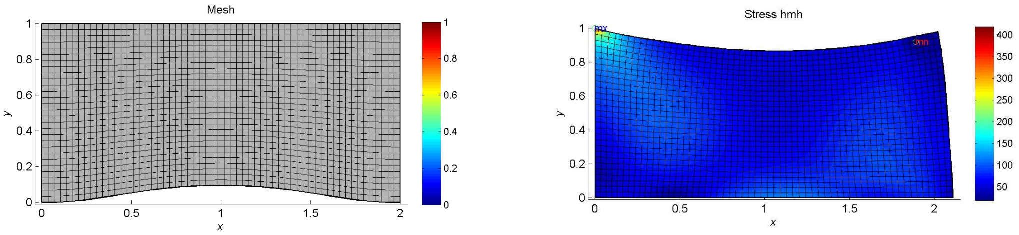
$$\left.\begin{array}{rl} \text{minimize} & |\boldsymbol{\lambda}(\boldsymbol{\alpha})|_6^6, \\ \text{subj. to} & \boldsymbol{\alpha} \in U. \end{array}\right\}$$

Note that if $\boldsymbol{\alpha}^* \in U$ is such that $\boldsymbol{u}_\nu(\boldsymbol{\alpha}^*) < -\boldsymbol{\alpha}^*$, i.e., there is no contact between the deformed body and the obstacle, then by complementarity $\boldsymbol{\lambda}(\boldsymbol{\alpha}^*) = \mathbf{0}$ and hence $\boldsymbol{\alpha}^*$ solves the above shape optimization problem. In order to avoid such "trivial" cases, the volume constraint in U has to be imposed with a sufficiently large lower bound. Keeping the material parameters and forces equal to the ones used in Example 1, the constants in the definition of U are changed to $C_0 = 0.75$, $C_1 = 3$, $C_3 = 10$, $C_{21} = 1.8$ and $C_{22} = 2$ (essentially, no upper bound).

As in the previous example, Figure 4.6 shows the initial design before and after deformation; in Figure 4.7 the same situation is depicted in case of the optimal shape $\Omega(\boldsymbol{\alpha}_{\mathrm{opt}})$ as computed by the BT algorithm. During minimization the value of the cost functional was reduced by one order of magnitude from $6.3 \cdot 10^5$ to $7.3 \cdot 10^4$ in 140 iterations.

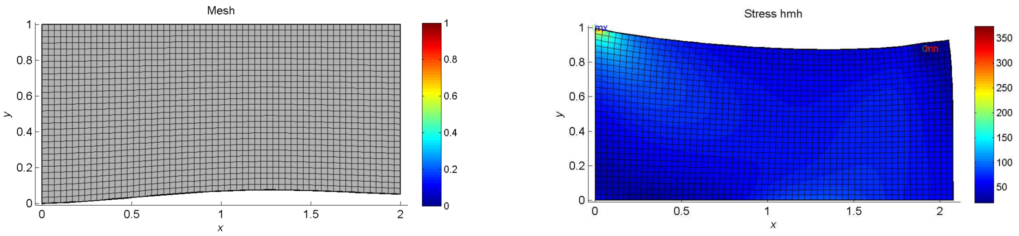
Example 3

The previous example related to the important technical issue of minimizing wear and fatigue by avoiding concentrations and peaks of contact stresses. In the case of *frictionless* contact problems it was shown in [25] that the aforementioned



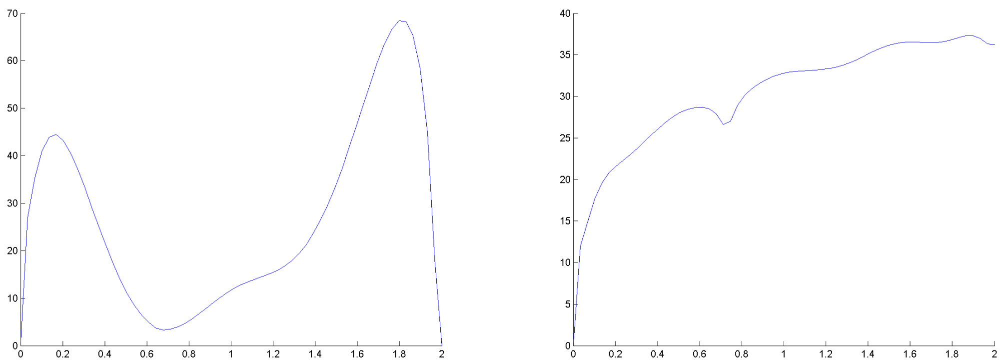
(a) Before deformation (b) After deformation

Figure 4.6: Example 2; initial design.



(a) Before deformation (b) After deformation

Figure 4.7: Example 2; optimal design.



(a) Initial design (b) Optimal design

Figure 4.8: Example 2; normal stresses.

effect is achieved by minimizing the total *potential energy* functional

$$\mathcal{E}(\boldsymbol{\alpha}) := E(\boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{\alpha})) = \frac{1}{2}\boldsymbol{u}^T(\boldsymbol{\alpha})\mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u}(\boldsymbol{\alpha}) - \boldsymbol{L}^T(\boldsymbol{\alpha})\boldsymbol{u}(\boldsymbol{\alpha}). \tag{4.20}$$

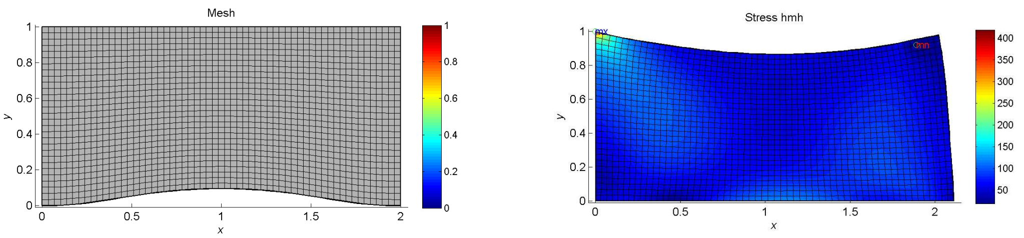
Importantly, it can be shown that in this case $\mathcal{E}(\boldsymbol{\alpha})$ is *continuously* differentiable. This has a considerable impact on the minimization algorithm, among others.

Unfortunately, solutions to contact problems with Coulomb friction cannot be described as a minimizer of some quadratic functional, like $E(\boldsymbol{\alpha}, \cdot)$. Nevertheless, we may still ask what do we get by minimizing the elastic energy, whether the optimal shape has similar properties as in the frictionless case. To this end we define the shape optimization problem:

$$\left.\begin{array}{rl} \text{minimize} & \mathcal{E}(\boldsymbol{\alpha}), \\ \text{subj. to} & \boldsymbol{\alpha} \in U, \end{array}\right\}$$

but in the definition (4.20) of the cost functional $\mathcal{E}$ the function $\boldsymbol{u}(\boldsymbol{\alpha})$ now stands for the first component of the solution to $(M^C(\boldsymbol{\alpha}))$. All parameters (coefficient of friction, material parameters, forces, constants in the definition of U, initial design, etc.) are the same as in the previous example.

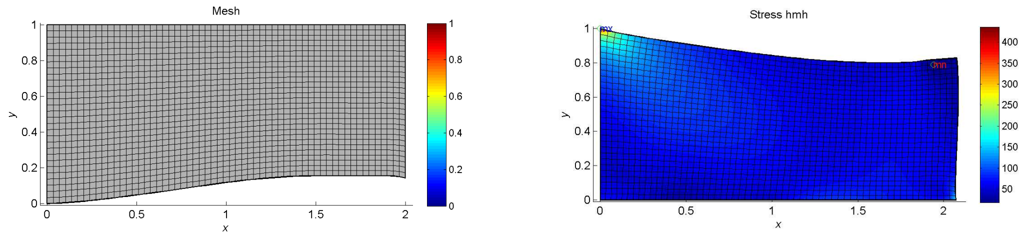
This time the BT solver took 48 iterations to converge, yielding a decrease in the cost functional from the initial value -7.69 to -10.88. Comparing the obtained optimal shape $\tilde{\boldsymbol{\alpha}}_{\text{opt}}$ (see Figure 4.10) with Figure 4.7, the resemblence is significant. In particular, the distribution of the normal contact stress (see right-hand side picture in Figure 4.11) is "almost constant" along $\Gamma_C(\tilde{\boldsymbol{\alpha}}_{\text{opt}})$, with around the same value as on the right of Figure 4.8—except for the node, where the contact and Neumann boundary conditions meet.



(a) Before deformation

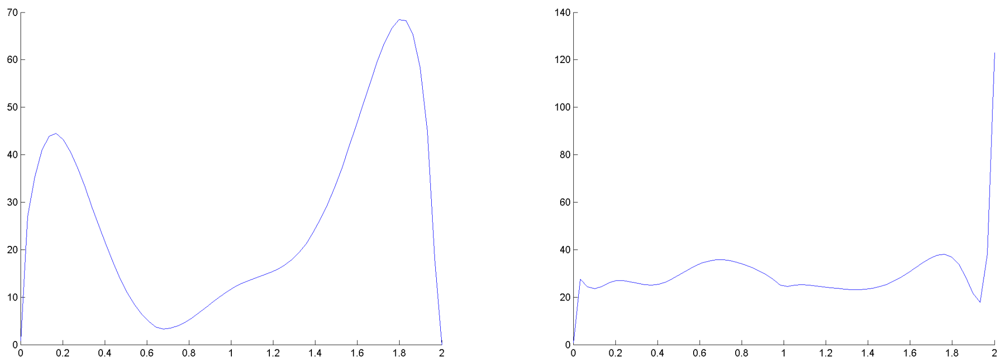
(b) After deformation

Figure 4.9: Example 3; initial design.



(a) Before deformation (b) After deformation

Figure 4.10: Example 3; optimal design.



(a) Initial design (b) Optimal design

Figure 4.11: Example 3; normal stresses.

Conclusion

Summary

In the thesis we address two separate, but related problems, namely, shape optimization in contact problems with two different models of friction: the Tresca and Coulomb laws of friction. In both cases we assume that the friction coefficient may depend on the solution. In order to see the commonalities and differences in the analysis of these problems side by side, the thesis is divided into four logical units as follows.

Chapter 1 concerns frictional contact problems in general. We start with the Signorini problem with Tresca friction—not just because of its simplicity, but it will also serve as a common basis for the analysis of the two friction laws mentioned above and of our interest. For all three frictional contact problems we mention their weak forms, discretize them and derive their algebraic counterparts. In addition to the usual primal formulation of the variational inequalities representing the weak form of our contact problems we also give their so-called mixed formulation. These involve the normal contact stresses as Lagrange multipliers—an important physical quantity which is of interest in many applications, not to mention the Coulomb friction model itself. Since our aim is to solve the shape optimization problems numerically, we focus on the algebraic state problems and give appropriate conditions on the friction coefficient ensuring their unique solvability. Moreover, these conditions do not depend on the geometry, as noted in Chapter 1.

The main part of the thesis is composed of chapters 2 and 3, in which we investigate the shape optimization problems linked to the state problems described above. These are treated on the algebraic level only and take the form of an MPEC. Our goal is not to analyze the MPECs for one particular cost functional, but rather the ability to choose cost functionals from a broad family. To this end we specify an admissible set for the shape parameter in the beginning and show that the shape optimization problems attain a solution for any "reasonable" cost functional provided the friction coefficient is regular enough. Obviously, these conditions differ for the Tresca and Coulomb models of friction, but in both cases lead to unique solvability of the respective state problems. Therefore, it is natural to approach the numerical solution of the MPECs via the ImP method. However, in order to apply subgradient methods to the minimization of the resulting NLP, one has to be able to compute (Clarke's) subgradients of the nonsmooth, nonconvex, implicitly defined control-to-state mappings. This matter is addressed in the respective sections devoted to sensitivity analysis. Here we make extensive use of modern tools from variational analysis, in particular the generalized differential calculus of B. Mordukhovich.

Chapter 4 is devoted to the numerical solution of the shape optimization problems in line with the ideas outlined above. We brifly sketch the BT algorithm user for the minimization of both NLP problems resulting from the ImP approach. Next, the adjoint generalized equations, derived in the previous chapters for the computation of subgradients, are revisited and their solution explained in more detail. Finally, the theoretical results are demonstrated by three examples: one using the Tresca model of friction and two involving the Coulomb model. In each case we use a different cost functional, demonstrating various features of contact shape optimization problems.

For the convenience of the uninitiated reader we have alse included an appendix, in which we gather basic definitions from the theory of nonsmooth and variational analysis. In particular, we discuss various notions from Clarke's and Mordukhovich's calculus and their relationship, but only to the extent needed in the thesis.

Outlook

Finally, let us outline some directions and areas for future research, improving on the results obtained in the present thesis.

A straightforward follow-up on the thesis would be the generalization of the state problem to three space dimensions. The 3D Signorini problem with Tresca friction involving a solution-dependent coefficient of friction was analyzed in [30] and the 3D Signorini problem with Coulomb friction involving a solution-dependent friction coefficient in [31]. The results of these papers are comparable to those in 2D, in particular the discretized contact problems are uniquely solvable provided the friction coefficient satisfies some regularity and smallness assumptions. For shape optimization it is essential that these assumptions do not depend on the geometry of the underlying domain (if chosen from a suitable family of admissible ones). The analysis presented in Chapter 2 and 3 seems to be fairly straightforward to implement in the three-dimensional setting up to the AGE. The only difference is in the computation of the coderivative of the multifunction from the state GE, that is indespensable for the numerical solution of the shape optimization problem as presented here. At this point, ideas from the thesis and [4] could possibly be combined and refined in order to derive an expression that may already be evaluated in computer code.

Notice that throughout the thesis we silently assumed that the cost functional depends only on the contact displacements $\boldsymbol{u}_\tau$, $\boldsymbol{u}_\nu$ and the normal contact stress $\boldsymbol{\lambda}$. In some applications, however, the tangential contact stress (related to the friction force) might be subject to optimization, as well. To deal with this situation, two possible solutions come immediately into ones mind. We shall sketch them briefly. In Section 4.2.2 we have already seen that the tangential contact stress, let us denote it by $\boldsymbol{\lambda}_\tau$, may be incorporated into the state problem in the form of another Lagrange multiplier—as it was the case with the normal contact stress, denoted by $\boldsymbol{\lambda}_\nu$ hereafter. From the first equation in (4.7) we can express

$$\mathbb{T}^T \boldsymbol{\lambda}_\tau = \boldsymbol{L}(\boldsymbol{\alpha}) - \mathbb{A}(\boldsymbol{\alpha})\boldsymbol{u} - \mathbb{N}^T \boldsymbol{\lambda}_\nu. \tag{4.21}$$

Therefore, one possibility to calculate the sensitivity of $\boldsymbol{\lambda}_\tau$ with respect to $\boldsymbol{\alpha}$ is to apply the sum rule on the right-hand side of (4.21) and combine it with the results obtained in Chapter 2 and 3. Since the qualification condition ensuring equality in the nonsmooth sum rule might be difficult or impossible to prove, one may consider an alternative way as outlined below.

Eliminating from (4.7) the state variables which correspond to the "internal" nodes of the triangulation (as it was done in Chapter 1 to get (1.34)), one arrives at the following GE:

$$\mathbf{0} \in \begin{bmatrix} \mathbb{A}_{\tau\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\tau\nu}(\boldsymbol{\alpha}) & \mathbb{E} & \mathbf{0} \\ \mathbb{A}_{\nu\tau}(\boldsymbol{\alpha}) & \mathbb{A}_{\nu\nu}(\boldsymbol{\alpha}) & \mathbf{0} & -\mathbb{E} \\ -\mathbb{E} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{E} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_\tau \\ \boldsymbol{u}_\nu \\ \boldsymbol{\lambda}_\tau \\ \boldsymbol{\lambda}_\nu \end{bmatrix} - \begin{bmatrix} \boldsymbol{L}_\tau(\boldsymbol{\alpha}) \\ \boldsymbol{L}_\nu(\boldsymbol{\alpha}) \\ \mathbf{0} \\ -\boldsymbol{\alpha} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ N_{\Lambda_\tau(\boldsymbol{\alpha},\boldsymbol{\varphi},\boldsymbol{g})}(\boldsymbol{\lambda}_\tau) \\ N_{\mathbb{R}^p_+}(\boldsymbol{\lambda}_\nu) \end{bmatrix}. \tag{4.22}$$

To derive the GEs corresponding to the contact problems investigated in the thesis, it is sufficient to apply in (4.22) the respective fixed-point properties, see Definition 6 and the discussion below Definition 9. In both cases the resulting GEs take the form assumed in [37] and thus sensitivity analysis may be carried out using the results of [37]. However, there is a substantial difference: in order for the basic assumption (3.1) in [37] to hold, $\mathfrak{F}$ needs to be *twice continuously differentiable*, whereas the analysis in Chapter 2 and 3 required basically no further smoothness of $\mathfrak{F}$ besides Lipschitz continuity. This apparent discrepancy might be also interesting to investigate.

Appendix A

Elements of variational analysis

When working with variational inequalities and optimal shape design problems, one inevitably comes across functions and mappings that are not (continuously) differentiable everywhere in their respective domain of definition. In order to investigate their differential properties, new tools had to be introduced that extend the classical calculus to functions which are not necessarily smooth or even single-valued.

The purpose of this chapter is to collect basic notions from nonsmooth and variational analysis that are extensively used in the last three chapters of the present thesis. The first section is devoted to the classical definition of Clarke's calculus for locally Lipschitz functions, in the second section we give basic definitions from the generalized differential calculus of Mordukhovich and present the relationship of the two thoeries.

A.1 Clarke calculus

A most prominent tool to treat functions that are (locally) Lipschitz, but not necessarily differentiable or convex, is the subdifferential calculus developed by Clarke [7]. Let us start with the definition of Lipschitz continuity of a function defined on a finite-dimensional Euclidean space (we shall work in finite dimensions throughout our presentation).

Definition 11 (Lipschitz continuity)**.** Let $n, m \in \mathbb{N}$ and $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$. We say, that $\boldsymbol{F}$ is

(i) *Lipschitz on* $\emptyset \neq M \subset \mathbb{R}^n$ iff there exists a constant $K \geq 0$ such that

$$\|\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})\|_m \leq K\|\boldsymbol{x} - \boldsymbol{y}\|_n \quad \forall \boldsymbol{x}, \boldsymbol{y} \in M; \tag{A.1}$$

(ii) *Lipschitz around* $\boldsymbol{x}$ iff there exists a neighbourhood $\mathcal{U}$ of $\boldsymbol{x}$ such that $\boldsymbol{F}$ is Lipschitz on $\mathcal{U}$;

(iii) *locally Lipschitz* iff $\boldsymbol{F}$ is Lipschitz around each $\boldsymbol{x}$ from its domain of definition.

Let $\boldsymbol{F}$ be Lipschitz around $\boldsymbol{x}$. Then it is evident from (A.1) that $\boldsymbol{F}$ is also continuous at $\boldsymbol{x}$ and the set $\left\{\frac{1}{t}(\boldsymbol{F}(\boldsymbol{y} + t\boldsymbol{v}) - \boldsymbol{F}(\boldsymbol{y})) \mid |t| \text{ sufficiently small} \right\}$ is uniformly bounded with respect to $\boldsymbol{v} \in \mathbb{R}^n$, $\|\boldsymbol{v}\|_n = 1$, and $\boldsymbol{y}$ sufficiently close to $\boldsymbol{x}$. However, $\boldsymbol{F}$ need not be directionally differentiable at $\boldsymbol{x}$. Nevertheless, a fundamental property of Lipschitzian functions, proved by Rademacher [45], is the fact that the set of such points is small (in a sense that it has zero Lebesgue measure).

Lemma 11 (Rademacher)**.** *Let $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz. Then*

$$\Omega_F := \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{F} \textit{ is not Fréchet differentiable at } \boldsymbol{x}\}$$

has Lebesgue measure 0.

One possible way to develop calculus for Lipschitzian functions is to give a suitable definition of directional derivatives and (sub)gradients—this approach is followed below. In the next section, where we introduce the Mordukhovich generalized differential calculus, we shall give an equivalent formulation of these notions from a variational geometry point of view, i.e., based on tangential and normal cones to the epigraph of a function.

Definition 12 (Clarke's generalized directional derivative)**.** Let $\boldsymbol{x}, \boldsymbol{v} \in \mathbb{R}^n$ be arbitrary and $f : \mathbb{R}^n \to \mathbb{R}$ Lipschitz around $\boldsymbol{x}$. The value

$$f^0(\boldsymbol{x}; \boldsymbol{v}) := \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0_+}} \frac{f(\boldsymbol{y} + t\boldsymbol{v}) - f(\boldsymbol{y})}{t}$$

is then called Clarke's *generalized directional derivative* of f at $\boldsymbol{x}$ in direction $\boldsymbol{v}$.

Definition 13 (Clarke's generalized (sub)gradient)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz around $\boldsymbol{x}$. Then the set

$$\bar{\partial} f(\boldsymbol{x}) := \{\boldsymbol{\xi} \in \mathbb{R}^n \mid f^0(\boldsymbol{x}; \boldsymbol{v}) \geq \langle \boldsymbol{\xi}, \boldsymbol{v} \rangle_n \; \forall \boldsymbol{v} \in \mathbb{R}^n\}$$

is called the *Clarke subdifferential* of f at $\boldsymbol{x}$ and its elements are Clarke's *generalized gradients* (or *Clarke's subgradients*).

It turns out that for a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ the following useful relation holds between its generalized directional derivative and gradients (see [7]):

$$f^0(\boldsymbol{x}; \boldsymbol{v}) = \max\{\langle \boldsymbol{\xi}, \boldsymbol{v} \rangle_n \mid \boldsymbol{\xi} \in \bar{\partial} f(\boldsymbol{x})\}.$$

If f happens to be continuously differentiable around $\boldsymbol{x}$, then $f^0(\boldsymbol{x}; \cdot)$ and $\bar{\partial} f(\boldsymbol{x})$ coincide with the classical directional derivative $f'(\boldsymbol{x}; \cdot)$ and gradient $\nabla f(\boldsymbol{x})$, respectively.

Due to Rademacher's lemma one may express Clarke's subdifferential in the following equivalent form—we refer to [7] for its proof.

Theorem 22. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz around $\boldsymbol{x} \in \mathbb{R}^n$. Then*

$$\bar{\partial} f(\boldsymbol{x}) = \operatorname{conv} \left\{ \lim_{i \to \infty} \nabla f(\boldsymbol{x}^{(i)}) \mid \boldsymbol{x}^{(i)} \to \boldsymbol{x}, \; \boldsymbol{x}^{(i)} \notin \Omega_f \right\}.$$

On the basis of the above theorem one may generalize the notion of Clarke's subdifferential to vector-valued Lipschitzian mappings $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$.

Definition 14 (Clarke's generalized Jacobian)**.** Let $m, n \in \mathbb{N}$ and $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz around $\boldsymbol{x} \in \mathbb{R}^n$. Then the set

$$\bar{\partial}\boldsymbol{F}(\boldsymbol{x}) := \left\{ \lim_{i\to\infty} \nabla \boldsymbol{F}(\boldsymbol{x}^{(i)}) \mid \boldsymbol{x}^{(i)} \to \boldsymbol{x},\ \boldsymbol{x}^{(i)} \notin \Omega_F \right\} \subset \mathbb{R}^{m\times n}$$

is called *Clarke's generalized Jacobian* of $\boldsymbol{F}$ at $\boldsymbol{x}$.

It can be immediately seen that $\bar{\partial}\boldsymbol{F}(\boldsymbol{x})$ is nonempty and compact, whenever the assumptions of Definition 14 are met. In addition, the generalized gradients and Jacobians introduced in Definition 13 and Definition 14, resp., enjoy rather rich calculus rules for computing generalized gradients or Jacobians of sums or compositions of locally Lipschitz mappings, cf. [7]. These rules are usually in the form of set inclusions, provided some additional qualification conditions are met. In case of additional smoothness and regularity assumptions these inclusions turn into equalities.

One smoothness condition that ensures directional differentiability of a locally Lipschitz mapping, but is weaker than Fréchet (or continuous) differentiability is that of *semismoothness*. It was first introduced by Mifflin [32] for Lipschitzian functions $f : \mathbb{R}^n \to \mathbb{R}$ and later generalized to vector-valued mappings by Qi and Sun [44].

Definition 15 (semismoothness)**.** Let $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz around $\boldsymbol{x} \in \mathbb{R}^n$. We say that $\boldsymbol{F}$ is

(i) *semismooth* at $\boldsymbol{x}$ iff the limit

$$\lim_{\substack{\mathbb{V}\in\bar{\partial}\boldsymbol{F}(\boldsymbol{x}+t\boldsymbol{v}') \\ \boldsymbol{v}'\to\boldsymbol{v},\ t\to 0_+}} \{\mathbb{V}\boldsymbol{v}'\}$$

exists for all $\boldsymbol{v} \in \mathbb{R}^n$;

(ii) *weakly semismooth* at $\boldsymbol{x}$ iff the limit

$$\lim_{\substack{\mathbb{V}\in\bar{\partial}\boldsymbol{F}(\boldsymbol{x}+t\boldsymbol{v}) \\ t\to 0_+}} \{\mathbb{V}\boldsymbol{v}\}$$

exists for all $\boldsymbol{v} \in \mathbb{R}^n$.

It is clear, that semismoothness implies weak semismoothness. Moreover, it holds that if $\boldsymbol{F}$ is weakly semismooth at $\boldsymbol{x}$, than it is also directionally differentiable at $\boldsymbol{x}$ and

$$\boldsymbol{F}'(\boldsymbol{x};\boldsymbol{v}) = \lim_{\substack{\mathbb{V}\in\bar{\partial}\boldsymbol{F}(\boldsymbol{x}+t\boldsymbol{v}) \\ t\to 0_+}} \{\mathbb{V}\boldsymbol{v}\}$$

for every $\boldsymbol{v} \in \mathbb{R}^n$ (cf. [44, Proposition 2.1]). Smooth, piecewise smooth, or convex functions are all examples of semismooth functions.

A.2 Mordukhovich calculus

A.2.1 Multifunctions

We start by collecting the most basic notions from set-valued analysis that are going to be used in the thesis. For a more thorough presentation of the topic we kindly refer to e.g. [2].

Let us recall that by a *set-valued mapping* (or simply *multifunction*) $F : X \rightrightarrows Y$ we mean a function $F : X \to 2^Y$, i.e., $F(x) \subset Y$ for each $x \in X$. The sets

$$\begin{aligned} \operatorname{Dom} F &:= \{x \in X \mid F(x) \neq \emptyset\}, \\ \operatorname{Gr} F &:= \{(x,y) \in X \times Y \mid x \in X,\ y \in F(x)\}, \end{aligned}$$

are called the *domain* and *graph* of F, respectively. We use the common term *closed multifunction* if $\operatorname{Gr} F$ is closed in the product topology of $X \times Y$.

In the sequel we shall restrict our presentation to the finite dimensional case, i.e., when $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ for some $n, m \in \mathbb{N}$.

Definition 16 (Kuratowski-Painlevé outer/inner limit of sets)**.** Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a multifunction and $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ arbitrary. Then the sets

$$\begin{aligned} \operatorname*{Lim\,sup}_{\boldsymbol{x}\to\bar{\boldsymbol{x}}} F(\boldsymbol{x}) &:= \{\boldsymbol{y} \in \mathbb{R}^m \mid \exists \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}\ \exists \boldsymbol{y}^{(i)} \to \boldsymbol{y} :\ \boldsymbol{y}^{(i)} \in F(\boldsymbol{x}^{(i)})\}, \\ \operatorname*{Lim\,inf}_{\boldsymbol{x}\to\bar{\boldsymbol{x}}} F(\boldsymbol{x}) &:= \{\boldsymbol{y} \in \mathbb{R}^m \mid \forall \boldsymbol{x}^{(i)} \to \bar{\boldsymbol{x}}\ \exists \boldsymbol{y}^{(i)} \to \boldsymbol{y} :\ \boldsymbol{y}^{(i)} \in F(\boldsymbol{x}^{(i)})\} \end{aligned}$$

are called the *Kuratowski-Painlevé outer* and *inner limit* of F at $\bar{\boldsymbol{x}}$, respectively.

Several Lipschitz-like properties may be defined for multifunctions. A direct generalization of local Lipschitz continuity of single-valued functions as introduced in Definition 11 is the so-called *Aubin property* (originally the term *pseudo-Lipschitzian* property was used by Aubin [1]).

Definition 17 (Aubin property)**.** A multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to have the *Aubin property around* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} F$ iff there exist neighbourhoods $\mathcal{U}$ of $\bar{\boldsymbol{x}}$ and $\mathcal{V}$ of $\bar{\boldsymbol{y}}$ and a constant $K \geq 0$ so that

$$F(\boldsymbol{x}) \cap \mathcal{V} \subset F(\boldsymbol{x}') + K\|\boldsymbol{x} - \boldsymbol{x}'\|_n \overline{\mathbb{B}}_m \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{U}. \tag{A.2}$$

It can be seen that if F happens to be single-valued around $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$, $\bar{\boldsymbol{y}} = F(\bar{\boldsymbol{x}})$, the above definition reduces to that of Lipschitz continuity around $\bar{\boldsymbol{x}}$.

By fixing $\boldsymbol{x}' \equiv \bar{\boldsymbol{x}}$ in (A.2) we arrive at the weaker property called *calmness.* It was originally introduced in [53, Definition 2.8] under the term *pseudo upper-Lipschitz continuity.*

Definition 18 (calmness)**.** A multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *calm around* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} F$ iff there exist neighbourhoods $\mathcal{U}$ of $\bar{\boldsymbol{x}}$ and $\mathcal{V}$ of $\bar{\boldsymbol{y}}$ and a constant $K \geq 0$ such that

$$F(\boldsymbol{x}) \cap \mathcal{V} \subset F(\bar{\boldsymbol{x}}) + K\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_n \overline{\mathbb{B}}_m \quad \forall \boldsymbol{x} \in \mathcal{U}.$$

A.2.2 Generalized differentiation

Definition 19 (contingent cone)**.** Let $\emptyset \neq A \subset \mathbb{R}^n$ and $\bar{\boldsymbol{x}} \in \overline{A}$ be arbitrary. Then the set

$$T_A(\bar{\boldsymbol{x}}) := \operatorname*{Lim\,sup}_{\lambda\to 0_+} \frac{A - \bar{\boldsymbol{x}}}{\lambda} \tag{A.3}$$

is called the *contingent cone* (or *Bouligand tangent cone*) to A at $\bar{\boldsymbol{x}}$.

Definition 20 (regular and limiting normal cone)**.** Let $\emptyset \neq A \subset \mathbb{R}^n$ and $\bar{\boldsymbol{x}} \in \overline{A}$ be arbitrary. The *regular (Fréchet) normal cone* to A at $\bar{\boldsymbol{x}}$ is defined as

$$\widehat{N}_A(\bar{\boldsymbol{x}}) := \Big\{ \boldsymbol{x}^* \in \mathbb{R}^n \;\Big|\; \limsup_{\boldsymbol{x} \xrightarrow{A} \bar{\boldsymbol{x}}} \frac{\langle \boldsymbol{x}^*, \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle_n}{\| \boldsymbol{x} - \bar{\boldsymbol{x}} \|_n} \leq 0 \Big\}. \tag{A.4}$$

For $\bar{\boldsymbol{x}} \notin A$ one sets $\widehat{N}_A(\bar{\boldsymbol{x}}) := \emptyset$. The *limiting (Mordukhovich) normal cone* to A at $\bar{\boldsymbol{x}}$ is then defined as

$$N_A(\bar{\boldsymbol{x}}) := \operatorname*{Lim\,sup}_{\boldsymbol{x} \xrightarrow{A} \bar{\boldsymbol{x}}} \widehat{N}_A(\boldsymbol{x}). \tag{A.5}$$

All three sets introduced in (A.3)–(A.5) are indeed closed cones with vertex at $\mathbf{0}$ and–in the assumed finite dimensional setting–the relation

$$\widehat{N}_A(\bar{\boldsymbol{x}}) = (T_A(\bar{\boldsymbol{x}}))^0 = \{ \boldsymbol{x}^* \in \mathbb{R}^n \mid \langle \boldsymbol{x}^*, \boldsymbol{v} \rangle_n \leq 0 \; \forall \boldsymbol{v} \in T_A(\bar{\boldsymbol{x}}) \} \tag{A.6}$$

holds true, where C^0 denotes the (negative) polar cone to C.

Remark 14. Due to (A.6) the regular normal cone $\widehat{N}_A(\bar{\boldsymbol{x}})$ is always *convex*, whereas $T_A(\bar{\boldsymbol{x}})$ and $N_A(\bar{\boldsymbol{x}})$ are in general *nonconvex*. This means that the limiting normal cone *cannot* be expressed as the dual to any tangent cone.

It can be immediately seen that the inclusion

$$\widehat{N}_A(\bar{\boldsymbol{x}}) \subset N_A(\bar{\boldsymbol{x}}) \tag{A.7}$$

holds for any nonempty $A \subset \mathbb{R}^n$ and $\bar{\boldsymbol{x}} \in \overline{A}$. If (A.7) holds with equality, we say that the set A is *normally regular* at $\bar{\boldsymbol{x}}$. E.g., if A is locally convex around $\bar{\boldsymbol{x}}$, it is automatically normally regular at this point.

Given an extended-real-valued function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$, its *epigraph* is the set

$$\operatorname{epi} \varphi := \{ (\boldsymbol{x}, y) \in \mathbb{R}^n \times \mathbb{R} \mid \boldsymbol{x} \in \mathbb{R}^n, \; y \geq \varphi(\boldsymbol{x}) \}.$$

On the basis of (A.4) and (A.5) one may define various subdifferentials of φ as suitable sets of normals to its epigraph.

Definition 21 (regular and limiting subdifferential)**.** Let $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{\boldsymbol{x}} \in \mathbb{R}^n$. Then

$$\widehat{\partial} \varphi(\bar{\boldsymbol{x}}) := \{ \boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -1) \in \widehat{N}_{\operatorname{epi} \varphi}(\bar{\boldsymbol{x}}, \varphi(\bar{\boldsymbol{x}})) \} \tag{A.8}$$

is called the *regular subdifferential* of φ at $\bar{\boldsymbol{x}}$, whereas

$$\partial \varphi(\bar{\boldsymbol{x}}) := \{ \boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -1) \in N_{\operatorname{epi} \varphi}(\bar{\boldsymbol{x}}, \varphi(\bar{\boldsymbol{x}})) \} \tag{A.9}$$

stands for the *limiting subdifferential* of φ at $\bar{\boldsymbol{x}}$.

If φ is lower semicontinuous around $\bar{\boldsymbol{x}}$ (i.e., its epigraph is closed around $(\bar{\boldsymbol{x}}, \varphi(\bar{\boldsymbol{x}}))$), then the limiting subdifferential may be expressed as

$$\partial \varphi(\bar{\boldsymbol{x}}) = \operatorname*{Lim\,sup}_{\boldsymbol{x} \xrightarrow{\varphi} \bar{\boldsymbol{x}}} \widehat{\partial} \varphi(\boldsymbol{x}), \tag{A.10}$$

where $\boldsymbol{x} \xrightarrow{\varphi} \bar{\boldsymbol{x}}$ means that $\boldsymbol{x} \to \bar{\boldsymbol{x}}$ with $\varphi(\boldsymbol{x}) \to \varphi(\bar{\boldsymbol{x}})$. Of course, if φ is convex around $\bar{\boldsymbol{x}}$, then both $\widehat{\partial} \varphi(\bar{\boldsymbol{x}})$ and $\partial \varphi(\bar{\boldsymbol{x}})$ are equal to the classical convex subdifferential. In case φ is strictly differentiable at $\bar{\boldsymbol{x}}$, then $\widehat{\partial} \varphi(\bar{\boldsymbol{x}}) = \partial \varphi(\bar{\boldsymbol{x}}) = \{ \nabla \varphi(\bar{\boldsymbol{x}}) \}$, where $\nabla \varphi(\bar{\boldsymbol{x}})$ denotes the gradient of φ.

Considering the graph instead of the epigraph in (A.8) and (A.9), one may construct derivative-like objects for multifunctions, as well, called *coderivates*.

Definition 22 (regular and limiting coderivative)**.** Given a multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} F$, the multifunction from $\mathbb{R}^m$ into subsets of $\mathbb{R}^n$ defined by

$$\widehat{D}^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) := \{\boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in \widehat{N}_{\operatorname{Gr} F}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})\} \quad \forall \boldsymbol{y}^* \in \mathbb{R}^m,$$

is called the *regular coderivative* of F at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ in direction $\boldsymbol{y}^*$, whereas

$$D^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) := \{\boldsymbol{x}^* \in \mathbb{R}^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in N_{\operatorname{Gr} F}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})\} \quad \forall \boldsymbol{y}^* \in \mathbb{R}^m$$

denotes the *(limiting) coderivative* of F.

Since the normal cones (A.4) and (A.5) are pointed (contain the null vector), both coderivatives are positively homogeneous closed multifunctions for each $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} F$. In addition, they reduce to the adjoint Jacobian

$$\widehat{D}^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) = D^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\boldsymbol{y}^*) = \{(\nabla F(\bar{\boldsymbol{x}}))^T \boldsymbol{y}^*\}, \quad \boldsymbol{y}^* \in \mathbb{R}^m,$$

provided F is single-valued [1] and strictly differentiable at $\bar{\boldsymbol{x}}$. Moreover, the regular coderivative equals to the adjoint Jacobian $\widehat{D}^* F(\bar{\boldsymbol{x}})(\boldsymbol{y}^*) = \{(\nabla F(\bar{\boldsymbol{x}}))^T \boldsymbol{y}^*\}$, $\boldsymbol{y}^* \in \mathbb{R}^m$, whenever F is single-valued and Fréchet-differentiable at $\bar{\boldsymbol{x}}$; this does not hold for the limiting coderivative.

It has been found that the (limiting) coderivative may provide information about Lipschitzian behaviour of a closed multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ around $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in \operatorname{Gr} F$, since

$$F \text{ has the Aubin property around } (\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \iff D^* F(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})(\mathbf{0}) = \{\mathbf{0}\}, \quad \text{(A.11)}$$

see [36], [47]. The right hand side of the equivalence (A.11) is called the *Mordukhovich criterion*, proved by B. Mordukhovich in [34].

A.2.3 Application to Lipschitzian mappings

In this section we collect some facts concerning the application of generalized differentiation to single-valued and locally Lipschitz functions. In particular, we recall the relationship between the (limiting) coderivative and (limiting) subdifferential, and compare the (limiting) subdifferential with Clarke's subdifferential.

First of all, recall that the Aubin property reduces to local Lipschitz continuity in case of a single-valued mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, hence the Mordukhovich criterion (A.11) yields:

$$F \text{ is Lipschitz around } \bar{\boldsymbol{x}} \in \mathbb{R}^n \iff D^* F(\bar{\boldsymbol{x}})(\mathbf{0}) = \{\mathbf{0}\}. \quad \text{(A.12)}$$

The next result provides a convenient way for computing the coderivative of a locally Lipschitzian mapping via the limiting subdifferential.

Theorem 23 (scalarization formula)**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz around $\bar{\boldsymbol{x}} \in \mathbb{R}^n$. Then*

$$D^* F(\bar{\boldsymbol{x}})(\boldsymbol{y}^*) = \partial \langle \boldsymbol{y}^*, F \rangle (\bar{\boldsymbol{x}}) \quad \forall \boldsymbol{y}^* \in \mathbb{R}^m, \quad \text{(A.13)}$$

where $\langle \boldsymbol{y}^*, F \rangle : \boldsymbol{x} \mapsto \langle \boldsymbol{y}^*, F(\boldsymbol{x}) \rangle_m$, $\boldsymbol{x} \in \mathbb{R}^n$.

[1] For a single-valued mapping F we simply write $\widehat{D}^* F(\bar{\boldsymbol{x}})$ and $D^* F(\bar{\boldsymbol{x}})$, i.e., omit $\bar{\boldsymbol{y}} = F(\bar{\boldsymbol{x}})$ from the argument.

Proof. See e.g. [36, Theorem 3.28]. □

In the previous section we introduced Clarke's generalized derivative-like objects for Lipschitzian functions and mappings. In order to relate these notions to the limiting subdifferential and coderivative, let us rephrase them in terms of variational geometry.

For a given nonempty $A \subset \mathbb{R}^n$ and $\bar{\boldsymbol{x}} \in \overline{A}$, one may define the *Clarke tangent cone* to A at $\bar{\boldsymbol{x}}$ as

$$\bar{T}_A(\bar{\boldsymbol{x}}) := \operatorname*{Lim\,inf}_{\substack{\boldsymbol{x} \xrightarrow{A} \bar{\boldsymbol{x}}, \\ \lambda \to 0_+}} \frac{A - \boldsymbol{x}}{\lambda},$$

and the *Clarke normal cone* as its (negative) polar cone:

$$\bar{N}_A(\bar{\boldsymbol{x}}) := (\bar{T}_A(\bar{\boldsymbol{x}}))^0.$$

In particular, the Clarke normal cone is always convex. This way one has (cf. [47])

$$\bar{\partial}\varphi(\bar{\boldsymbol{x}}) = \{\boldsymbol{\xi} \in \mathbb{R}^n \mid (\boldsymbol{\xi}, -1) \in \bar{N}_{\operatorname{epi}\varphi}(\bar{\boldsymbol{x}}, \varphi(\bar{\boldsymbol{x}}))\} \tag{A.14}$$

for any $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ that is Lipschitz around $\bar{\boldsymbol{x}} \in \mathbb{R}^n$. In addition, (A.14) yields an extension of the Clarke subdifferential to more general functions, not necessarily Lipschitzian. Nevertheless, if φ is locally Lipschitz, then the following relation between its Clarke and Mordukhovich subdifferentials holds:

$$\bar{\partial}\varphi(\bar{\boldsymbol{x}}) = \operatorname{conv} \partial\varphi(\bar{\boldsymbol{x}}). \tag{A.15}$$

An analogous results holds true between the Clarke generalized Jacobian and the coderivative of a locally Lipschitz vector-valued mapping. This is formulated in a separate theorem below.

Theorem 24. *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz around $\bar{\boldsymbol{x}} \in \mathbb{R}^n$. Then*

$$(\bar{\partial}F(\bar{\boldsymbol{x}}))^T \boldsymbol{y}^* = \operatorname{conv} D^*F(\bar{\boldsymbol{x}})(\boldsymbol{y}^*) \quad \forall \boldsymbol{y}^* \in \mathbb{R}^m.$$

For the proofs of the respective claims in this section we refer to the monographs [36] and [47]. Finally, we conclude the chapter with an example that is intended to demonstrate all the above notions in a very simple situation.

Example 3. Let us define the set

$$A := \{(x, y) \in \mathbb{R}^2 \mid y \geq -|x|\},$$

i.e., A is the *epigraph* of $\varphi(x) = -|x|$, $x \in \mathbb{R}$. Clearly, φ is nonconvex, but Lipschitz with modulus 1. After some calculation, for the contingent and Clarke tangent cones at $(0, 0)$ one gets

$$T_A(0, 0) = A, \qquad \bar{T}_A(0, 0) = \{(h, k) \in \mathbb{R}^2 \mid k \geq |h|\},$$

respectively. Taking their negative polars yields the Fréchet and Clarke normal cones at $(0, 0)$:

$$\widehat{N}_A(0, 0) = \{(0, 0)\}, \qquad \bar{N}_A(0, 0) = \{(\xi, \eta) \in \mathbb{R}^2 \mid \eta \leq -|\xi|\}.$$

Therefore

$$\widehat{\partial}\varphi(0) = \emptyset \quad \text{and} \quad \bar{\partial}\varphi(0) = [-1, 1].$$

Since φ is smooth at all points except $x = 0$, we have $\widehat{\partial}\varphi(x) = \operatorname{sgn}(x)$ for $x \neq 0$ and (A.10) can be applied to determine the limiting subdifferential at 0:

$$\partial\varphi(0) = \{-1, 1\}. \tag{A.16}$$

When computing the coderivative mapping $D^*\varphi$ at 0, we first need to evaluate the normal cone to the *graph* of φ at $(0, 0)$:

$$B := \operatorname{Gr}\varphi = \{(x, y) \mid y = -|x|\}.$$

After some calculation one arrives at the expression

$$N_B(0, 0) = \{(x^*, y^*) \mid y^* \geq |x^*|\} \cup \{(x^*, y^*) \mid y^* = -|x^*|\},$$

where the first set on the right hand side equals $\widehat{N}_B(0, 0)$ and the second one represents limit points of $\widehat{N}_B(x, \varphi(x))$ as $x \to 0_+$ and $x \to 0_-$, that are not contained in the first set. From the definition of the coderivative we conclude

$$D^*\varphi(0)(y^*) = \begin{cases} \{-y^*, y^*\} & \text{if } y^* > 0, \\ \{0\} & \text{if } y^* = 0, \\ [y^*, -y^*] & \text{if } y^* < 0. \end{cases}$$

Notice that the case $y^* = 0$ is a consequence of the Mordukhovich criterion (A.12); the other cases may be computed employing the scalarization formula and positive homogeneity of the subdifferential mapping (i.e., $\partial(\alpha\varphi) = \alpha(\partial\varphi)$ for $\alpha \geq 0$):

$$D^*\varphi(0)(y^*) = \begin{cases} \big(\partial\varphi(0)\big)y^* = \{-1, 1\}y^*, & \text{if } y^* > 0 \\ \big(\partial(-\varphi)(0)\big)(-y^*) = [-1, 1](-y^*), & \text{if } y^* < 0. \end{cases}$$

In the first case we have used (A.16); in case $y^* < 0$ the function $-\varphi(x) = |x|$ is convex, therefore its subdifferential equals to the convex subdifferential.

Bibliography

[1] J.-P. Aubin, *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res. 9 (1984), 87–111.

[2] J.-P. Aubin, H. Frankowska, *Set-valued Analysis*, Birkhäuser, Boston, 1990.

[3] P. Beremlijski, J. Haslinger, M. Kočvara, J. Outrata, *Shape optimization in contact problems with Coulomb friction*, SIAM J. Optim. 13 (2002), 561–587.

[4] P. Beremlijski, J. Haslinger, M. Kočvara, R. Kučera, J. V. Outrata, *Shape optimization in three-dimensional contact problems with Coulomb friction*, SIAM J. Optim. 20(1) (2009), 416–444.

[5] P. Beremlijski, J. Haslinger, J. V. Outrata, R. Pathó, *Shape optimization in contact problems with Coulomb friction and a solution-dependent friction coefficient*, submitted to SIAM J. Control Optim.

[6] D. Chenais, *On the existence of a solution in a domain identification problem*, J. Math. Anal. App. 52 (1975), 189–219.

[7] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[8] A. L. Dontchev, W. W. Hager, *Implicit functions, Lipschitz maps and stability in optimization*, Math. Oper. Res. 19 (1994), 753–768.

[9] Z. Dostál, *Box constrained quadratic programming with proportioning and projections*, SIAM J. Optim. 7 (1997), 871–887.

[10] Z. Dostál, *Optimal Quadratic Programming Algorithms: With Applications to Variational Inequalities*, Springer Publishing Company, 2009.

[11] C. Eck, J. Jarušek, M. Krbec, *Unilateral Contact Problems: Variational Methods and Existence Theorems*, Chapman & Hall/CRC Press, Boca Raton, 2005.

[12] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[13] G. Fichera, *Sul problema elastostatico di Signorini con ambigue condizioni al contorno*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. 34 (1963), 138–142.

[14] G. Fichera, *Problemi elastostatici con vincoli unilaterali: il problema di Signorini con ambigue condizioni al contorno*, Atti Accad. Naz. Lincei, Mem., Cl. Sci. Fis. Mat. Nat., Sez. I, VIII. Ser. 7 (1964), 91–140.

[15] J. Haslinger, R. A. E. Mäkinen, *Introduction to Shape Optimization: Theory, Approximation, and Computation*, SIAM Advances in Design and Control Vol. 7, Philadelphia, 2003.

[16] J. Haslinger, P. Neittaanmäki, *Finite Element Approximation for Optimal Shape, Material and Topology Design*, 2nd ed., John Wiley & Sons, Chichester, 1996.

[17] J. Haslinger, J. V. Outrata, R. Pathó, *Shape optimization in 2D contact problems with given friction and a solution-dependent coefficient of friction*, Set-Valued Anal 20 (2012), 31–59.

[18] J. Haslinger, P. D. Panagiotopulos, *The reciprocal variational approach to the Signorini problem with friction. Approximation results*, Proc. R. Soc. Edinb. Sect. A 98 (1984), 365–383.

[19] J. Haslinger, O. Vlach, *Signorini problem with a solution dependent coefficient of friction (model with given friction): approximation and numerical realization*, Appl. Math. 50 (2005), 153–171.

[20] J. Haslinger, O. Vlach, *Approximation and numerical realization of 2D contact problems with Coulomb friction and a solution-dependent coefficient of friction*, J. Comput. Appl. Math. 197 (2006), 421–436.

[21] R. Henrion, A. Jourani, J. Outrata, *On the calmness of a class of multifunctions*, SIAM J. Optim. 13 (2002), 603–618.

[22] I. Hlaváček, J. Haslinger, J. Nečas, J. Lovíšek, *Solution of Variational Inequalities in Mechanics*, Springer-Verlag, New York, Berlin, 1988.

[23] N. Karmitsa, A. Bagirov, M. Mäkelä, *Comparing different nonsmooth minimization methods and software*, Optimization Methods and Software 27 (2012), 131–153.

[24] N. Kikuchi, J. T. Oden, *Contact Problems in Elasticity*, SIAM Press, Philadelphia, 1988.

[25] A. Klarbring, J. Haslinger, *On almost constant contact stress distribution by shape optimization*, Struct. Optimization 5 (1993), 213–216.

[26] D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization*, Springer, 2002.

[27] M. Kočvara, J. V. Outrata, *Optimization problems with equilibrium constraints and their numerical solution*, Math. Program. B 101 (2004), 119–150.

[28] T. Kozubek, A. Markopoulos, T. Brzobohatý, R. Kučera, V. Vondrák, Z. Dostál, *MatSol – MATLAB efficient solvers for problems in engineering*, `http://industry.it4i.cz/en/products/matsol/`.

[29] C. Lemaréchal, *Extensions diverses des méthods de gradient et applications*, Thè d'Etat, Paris, 1980.

[30] T. Ligurský, *Approximation and numerical realization of 3D contact problems with given friction and a coefficient of friction depending on the solution*, Master's thesis, Charles University in Prague, Faculty of Mathematics and Physics, 2007 (http://artax.karlin.mff.cuni.cz/~ligut2am/tl21.pdf).

[31] T. Ligurský, J. Haslinger, R. Kučera, *Approximation and numerical realization of 3D contact problems with Coulomb friction and a solution-dependent coefficient of friction*, Int. J. Numer. Meth. Engng 82 (2010), 1180–1206.

[32] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim. 15 (1977), 957–972.

[33] B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988 (in Russian).

[34] B. S. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, Theoretical Aspects of Industrial Design (D. A. Field and V. Komkov, eds.), SIAM Proc. Appl. Math., vol. 58, SIAM, Philadephia, 1992, 32–46.

[35] B. S. Mordukhovich, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl. 183 (1994), 250–288.

[36] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vols. 330 and 331, Springer-Verlag, Berlin-Heidelberg, 2006.

[37] B. S. Mordukhovich, J. V. Outrata, *Coderivative analysis of quasi-variational inequalities with applications to stability and optimization*, SIAM J Optim. 18 (2007), 389–412.

[38] M. M. Mäkelä, *Survey of bundle methods for nonsmooth optimization*, Optim. Methods Softw. 17(1) (2002), 1–29.

[39] M. M. Mäkelä, *Multiobjective proximal bundle method for nonconvex nonsmooth optimization: Fortran subroutine MPBNGC 2.0*, Reports of the Department of Mathematical Information Technology, Series B. Scientific Computing, B. 13/2003 University of Jyväskylä, Jyväskylä, 2003.

[40] M. M. Mäkelä, P. Neittaanmäki, *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*, World Scientific Publishing Co., Singapore, 1992.

[41] J. V. Outrata, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Math. Oper. Res. 24 (1999), 627–644.

[42] J. V. Outrata, J. Zowe, *A Newton method for a class of quasi-variational inequalities*, Comput. Optim. Appl. 4 (1995), 5–21.

[43] R. Pathó, *Shape optimization in contact problems with friction*, Master's Thesis, Charles University in Prague, Faculty of Mathematics and Physics, 2009 (http://www.karlin.mff.cuni.cz/~patho/dipl.pdf).

[44] L. Qi, J. Sun, *A nonsmooth version of Newton's method*, Math. Programming 58 (1993), 353–367.

[45] H. Rademacher, *Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale*, Mathematische Annalen 79 (1919), 340–359.

[46] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res. 5 (1980), 43–62.

[47] R. T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[48] S. Scholtes, *Introduction to Piecewise Differentiable Equations*, Habilitation Thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, Germany, 1994.

[49] C. H. Scholtz, *The Mechanics of Earthquakes and Faulting*, 2nd ed., Cambridge University Press, 2002.

[50] H. Schramm, J. Zowe, *A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results*, SIAM J. Optim. 2 (1992), 121–152.

[51] A. Signorini, *Questioni di elasticità non linearizzata e semilinearizzata*, Rend. Mat. Appl. 18 (1959), 95–139.

[52] J. Sokolowski, J.-P. Zolésio, *Introduction to Shape Optimization: Shape Sensitivity Analysis*, Springer-Verlag, Berlin, 1992.

[53] J. J. Ye, X. Y. Ye, *Necessary optimality conditions for optimization problems with variational inequality constraints*, Math. Oper. Res. 22 (1997), 977–997.

[54] Ph. Wolfe, *A method of conjugate gradients for minimizing nondifferentiable convex functions*, Math. Programming Stud. 3 (1975), 145–173.

[55] J. Zowe, H. Schramm, *Bundle Trust Methods: Fortran Codes for Nonsmooth Optimization. User's Guide*, Technical Report 259, Institute of Applied Mathematics, University of Erlangen, 2000.