PATH-FOLLOWING THE STATIC CONTACT PROBLEM WITH COULOMB FRICTION

J. Haslinger[1,2], V. Janovský[1], R. Kučera[2]

[1] Department of Numerical Mathematics, Charles University, Prague
Sokolovská 83, 186 75 Prague 8, Czech Republic
hasling@karlin.mff.cuni.cz, janovsky@karlin.mff.cuni.cz
[2] Department of Mathematics and Descriptive Geometry, VŠB-TU Ostrava
17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
radek.kucera@vsb.cz

Abstract

Consider contact problem with Coulomb friction on two planar domains. In order to find non-unique solutions we propose a new path following algorithm: Given a linear loading path we approximate the corresponding solution path. It consists of oriented piecewise linear branches connected by transition points. We developed a) predictor-corrector algorithm to follow oriented linear branches, b) branching and orientation indicators to detect transition points. The techniques incorporate semi-smooth Newton iterations and inactive/active set strategy on the contact zone.

1. Introduction

Consider deformable bodies in mutual contact. The relevant mathematical description consists in modeling both the non-penetration conditions and a friction law. The widely accepted Coulomb friction law represents a serious mathematical and numerical problem.

In particular, we consider 2D static contact problem with Coulomb friction. The problem is uniquely solvable, provided that the friction coefficient $\mathcal{F} > 0$ is sufficiently small, see [14, 6]. Since the seminal paper [14], no essential contribution was made concerning solvability of this problem for general data.

Obviously, engineers have always solved this important problem numerically, regardless unresolved theoretical issues. In a natural finite element (FEM) approximation, the discrete problem has always a solution, disregarding the size of $\mathcal{F}$, see [9, 8, 13]. Since the (discrete) problem is locally solvable, the idea was to apply the Implicit Function Theorem to follow the solution path, which was parameterized either by $\mathcal{F}$ or by a load increment. Nevertheless, lumped element models [11, 9, 13] indicate, that the particular solution points of interest should be those in which the Implicit Function Theorem *fails* to hold. They are *turning points* of the solution path. Actually, they are responsible for non-unique solvability of the problem.

The solution path is continuous, piecewise smooth, [8]. The classical numerical path following techniques, see e.g. [1], have to fail in principle. In [8], a special continuation algorithm was proposed to trace piecewise smooth solution curves. The algorithm was tested on lumped element models with just one or two points on the contact boundary, [12, 8].

In this paper, we present an improved continuation strategy and test it on a real FEM model. The outline is as follows: In Section 2, we define the state problem and its discretization. We recall the semi-smooth Newton method and apply it to the discrete state problem, see Section 3. The actual contribution is in Section 4, where a modified path following algorithm is presented. The substantial innovations consist in

1. application of *tangent continuation*, see [3], Algorithm 4.25,
2. introducing a robust *branching* and *orientation* indicator.

Note that due to material properties, the solution components are very uneven: The contact forces are within a range 10^6 N kg^{-1} while displacements are tiny.

2. State problem, FEM approximation

Let us consider two bodies Ω^1, Ω^2 in $\mathbb{R}^2$ with boundaries $\partial\Omega^k = \overline{\Gamma}_u^k \cup \overline{\Gamma}_p^k \cup \overline{\Gamma}_c^k$, $k = 1, 2$, see Figure 1. First, denote $\boldsymbol{u}^k$ the displacement field, $\boldsymbol{\sigma}(\boldsymbol{u}^k)$ the stress tensor, $\boldsymbol{f}^k$ the volume force, $\boldsymbol{p}^k$ the surface traction, $\boldsymbol{n}^k$ the outer normal vector to $\partial\Omega^k$, and $\lambda^k, \mu^k > 0$ material parameters. The state problem is defined by the Lamé equations in Ω^k, $k = 1, 2$,

$$
\begin{aligned}
-\mathrm{div}\,\boldsymbol{\sigma}(\boldsymbol{u}^k) &= \boldsymbol{f}^k, \\
\boldsymbol{\sigma}(\boldsymbol{u}^k) &= \lambda^k \mathrm{tr}(\boldsymbol{\epsilon}(\boldsymbol{u}^k))\boldsymbol{I} + 2\mu^k\boldsymbol{\epsilon}(\boldsymbol{u}^k), \\
\boldsymbol{\epsilon}(\boldsymbol{u}^k) &= \tfrac{1}{2}(\nabla\boldsymbol{u}^k + (\nabla\boldsymbol{u}^k)^\top),
\end{aligned}
$$
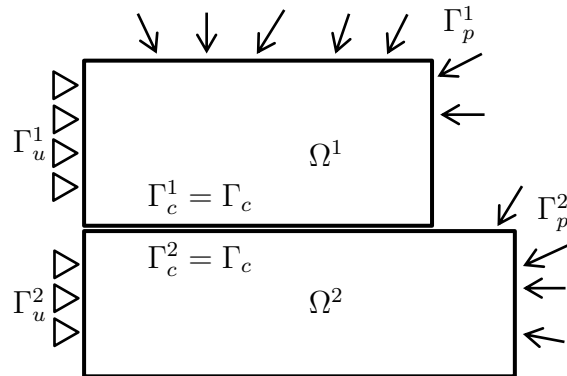

Figure 1: Geometry of the problem.

the Dirichlet and Neumann boundary conditions for $k = 1, 2$,

$$
\begin{aligned}
\boldsymbol{u}^k &= \boldsymbol{0} &&\text{on } \Gamma_u^k, \\
\boldsymbol{\sigma}(\boldsymbol{u}^k)\boldsymbol{n}^k &= \boldsymbol{p}^k &&\text{on } \Gamma_p^k,
\end{aligned}
$$

and by contact conditions on Γ_c:

- *unilateral contact law, Signorini problem:*

$$
u_\nu \leq 0, \ \sigma_\nu \leq 0, \ \sigma_\nu u_\nu = 0 \quad \text{on } \Gamma_c,
$$

  where $u_\nu = (\boldsymbol{u}^1 - \boldsymbol{u}^2)^\top \boldsymbol{n}$, $\sigma_\nu = \boldsymbol{n}^\top \boldsymbol{\sigma}(\boldsymbol{u}^1)\boldsymbol{n}$, and $\boldsymbol{n} = \boldsymbol{n}^1$,

- *transmission of contact stresses:*

$$
\boldsymbol{\sigma}(\boldsymbol{u}^1)\boldsymbol{n} = \boldsymbol{\sigma}(\boldsymbol{u}^2)\boldsymbol{n} \quad \text{on } \Gamma_c,
$$

- *the Coulomb friction law:*

$$
\begin{aligned}
|\sigma_t| &\leq -\mathcal{F}\sigma_\nu, \\
|\sigma_t| &< -\mathcal{F}\sigma_\nu &&\Rightarrow \ u_t = 0, \\
|\sigma_t| &= -\mathcal{F}\sigma_\nu &&\Rightarrow \ \exists c_t \geq 0: \ u_t = -c_t\sigma_t,
\end{aligned}
$$

  where $u_t = (\boldsymbol{u}^1 - \boldsymbol{u}^2)^\top \boldsymbol{t}$, $\sigma_t = \boldsymbol{t}^\top \boldsymbol{\sigma}(\boldsymbol{u}^1)\boldsymbol{n}$, $\boldsymbol{t}$ is orthogonal to $\boldsymbol{n}$, and $\mathcal{F} > 0$ is the coefficient of friction.

After FEM approximation we get the following *primal-dual* discrete state problem:

$$
\boldsymbol{K}\boldsymbol{u} + \boldsymbol{N}^\top \boldsymbol{\lambda}_\nu + \boldsymbol{T}^\top \boldsymbol{\lambda}_t = \boldsymbol{f}, \tag{1}
$$

$$
\boldsymbol{N}\boldsymbol{u} \leq 0, \ \boldsymbol{\lambda}_\nu \geq 0, \ \boldsymbol{\lambda}_\nu^\top \boldsymbol{N}\boldsymbol{u} = 0, \tag{2}
$$

$$
\left.
\begin{aligned}
|\lambda_{t,i}| &\leq \mathcal{F}\lambda_{n,i}, \\
|\lambda_{t,i}| &< \mathcal{F}\lambda_{n,i} \Rightarrow (\boldsymbol{T}\boldsymbol{u})_i = \boldsymbol{0}, \\
|\lambda_{t,i}| &= \mathcal{F}\lambda_{n,i} \Rightarrow \exists\, c_{t,i} \geq 0: \ (\boldsymbol{T}\boldsymbol{u})_i = c_{t,i}\lambda_{t,i},
\end{aligned}
\right\} \quad i = 1, \ldots, m, \tag{3}
$$

where $(\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$. Here $\boldsymbol{u}$ approximates the displacement field, $\boldsymbol{\lambda}_\nu$ and $\boldsymbol{\lambda}_t$ approximate normal and tangential stress components along the contact boundary Γ_c, m is the number of contact nodes. Data of the model: $\boldsymbol{K} \in \mathbb{R}^{n \times n}$ is positive definite stiffness matrix, $\boldsymbol{N}, \boldsymbol{T} \in \mathbb{R}^{m \times n}$ are full rank matrices (the actions of distributed contact forces along normal and tangential directions), $\boldsymbol{f} \in \mathbb{R}^n$ are nodal forces.

Next, we formulate inequalities (2)–(3) as a set of nonlinear equations using suitable projectors, see e.g. [7]. Let $P_{\mathbb{R}_+} : \mathbb{R} \mapsto \mathbb{R}_+$, $P_{\mathbb{R}_+}(x) = \max\{0, x\}$, $x \in \mathbb{R}$, be the projection onto $\mathbb{R}_+$. Let us define $P_{\mathbb{R}_+^m} : \mathbb{R}^m \mapsto \mathbb{R}_+^m$ for $\boldsymbol{x} = (x_1, \ldots, x_m)^\top$ by

$$
P_{\mathbb{R}_+^m}(\boldsymbol{x}) = (P_{\mathbb{R}_+}(x_1), \ldots, P_{\mathbb{R}_+}(x_m))^\top.
$$

Let $P_{[-g,g]} : \mathbb{R} \mapsto [-g, g]$, $P_{[-g,g]}(x) = \max\{0, x+g\} - \max\{0, x-g\} - g$, $x \in \mathbb{R}$, be the projection onto the interval $[-g, g]$, $g \geq 0$. Let us define $P_{[-\boldsymbol{g},\boldsymbol{g}]} : \mathbb{R}^m \mapsto [-\boldsymbol{g}, \boldsymbol{g}]$, where $[-\boldsymbol{g},\boldsymbol{g}] = [-g_1, g_1] \times \cdots \times [-g_m, g_m]$, $\boldsymbol{g} = (g_1, \ldots, g_m)^\top$, $g_i \geq 0$, for $\boldsymbol{x} = (x_1, \ldots, x_m)^\top$ by

$$P_{[-\boldsymbol{g},\boldsymbol{g}]}(\boldsymbol{x}) = (P_{[-g_1,g_1]}(x_1), \ldots, P_{[-g_m,g_m]}(x_m))^\top.$$

The inequalities (2) and (3) can be equivalently written as

$$\boldsymbol{\lambda}_\nu - P_{\mathbb{R}_+^m}(\boldsymbol{\lambda}_\nu + \rho \boldsymbol{N}\boldsymbol{u}) = \boldsymbol{0} \quad \text{and} \quad \boldsymbol{\lambda}_t - P_{[-\mathcal{F}\boldsymbol{\lambda}_\nu, \mathcal{F}\boldsymbol{\lambda}_\nu]}(\boldsymbol{\lambda}_t + \rho \boldsymbol{T}\boldsymbol{u}) = \boldsymbol{0},$$

respectively, where $\rho > 0$ is arbitrary but fixed (e.g., $\rho = 1$). Therefore, solving (1)–(3) is equivalent to finding roots $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ of the equation

$$G(\boldsymbol{y}) \equiv \begin{pmatrix} \boldsymbol{K}\boldsymbol{u} + \boldsymbol{N}^\top \boldsymbol{\lambda}_\nu + \boldsymbol{T}^\top \boldsymbol{\lambda}_t \\ \boldsymbol{\lambda}_\nu - P_{\mathbb{R}_+^m}(\boldsymbol{\lambda}_\nu + \rho \boldsymbol{N}\boldsymbol{u}) \\ \boldsymbol{\lambda}_t - P_{[-\mathcal{F}\boldsymbol{\lambda}_\nu, \mathcal{F}\boldsymbol{\lambda}_\nu]}(\boldsymbol{\lambda}_t + \rho \boldsymbol{T}\boldsymbol{u}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \\ 0 \end{pmatrix}, \tag{4}$$

where $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$.

The mapping $G : \mathbb{R}^{n+2m} \mapsto \mathbb{R}^{n+2m}$ is continuous and piecewise smooth. In particular, it is *piecewise affine*, see e.g. [16] for the notion.

3. The semi-smooth Newton method

To solve (4), we apply the Newton iterations. Due to the nature of the mapping G, semi-smooth methods are applicable [2]. Let us also refer to [10], where this technique was used for solving the Signorini problem.

Let $\mathcal{M} = \{1, 2, \ldots, m\}$ be the set of all indices of contact points. Given $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, we define the *inactive* sets $\mathcal{I}_\nu = \mathcal{I}_\nu(\boldsymbol{y})$, $\mathcal{I}_t^+ = \mathcal{I}_t^+(\boldsymbol{y})$, $\mathcal{I}_t^- = \mathcal{I}_t^-(\boldsymbol{y})$ by

$$\mathcal{I}_\nu = \{i \in \mathcal{M} : \lambda_{\nu,i} + \rho(\mathbf{N}\mathbf{u})_i < 0\},$$

$$\mathcal{I}_t^+ = \{i \in \mathcal{M} : \lambda_{t,i} + \rho(\mathbf{T}\mathbf{u})_i - \mathcal{F}\lambda_{\nu,i} > 0\},$$

$$\mathcal{I}_t^- = \{i \in \mathcal{M} : \lambda_{t,i} + \rho(\mathbf{T}\mathbf{u})_i + \mathcal{F}\lambda_{\nu,i} > 0\},$$

and the *active* sets $\mathcal{A}_\nu = \mathcal{A}_\nu(\boldsymbol{y})$, $\mathcal{A}_t = \mathcal{A}_t(\boldsymbol{y})$ as their complements:

$$\mathcal{A}_\nu = \mathcal{M} \setminus \mathcal{I}_\nu, \quad \mathcal{A}_t = \mathcal{M} \setminus (\mathcal{I}_t^+ \cup \mathcal{I}_t^-).$$

Let us introduce the indicator matrix $\boldsymbol{D}_\mathcal{S} \in \mathbb{R}^{m \times m}$ of $\mathcal{S} \subset \mathcal{M}$ as follows:

$$\boldsymbol{D}_\mathcal{S} = diag(s_1, \ldots, s_m), \quad s_i = \begin{cases} 1, & i \in \mathcal{S}, \\ 0, & i \in \mathcal{M} \setminus \mathcal{S}. \end{cases}$$

We observe that

$$
G(\boldsymbol{y}) =
\begin{pmatrix}
\boldsymbol{Ku} + \boldsymbol{N}^\top \boldsymbol{\lambda}_\nu + \boldsymbol{T}^\top \boldsymbol{\lambda}_t \\
\boldsymbol{\lambda}_\nu - \boldsymbol{D}_{\mathcal{A}_\nu}(\boldsymbol{\lambda}_\nu + \rho \boldsymbol{N} \boldsymbol{u}) \\
\boldsymbol{\lambda}_t - \boldsymbol{D}_{\mathcal{A}_t}(\boldsymbol{\lambda}_t + \rho \boldsymbol{T} \boldsymbol{u}) - \boldsymbol{D}_{\mathcal{I}_t^+}\mathcal{F}\boldsymbol{\lambda}_\nu + \boldsymbol{D}_{\mathcal{I}_t^-}\mathcal{F}\boldsymbol{\lambda}_\nu
\end{pmatrix}
= J(\boldsymbol{y})\,\boldsymbol{y}\,,
$$

where

$$
J(\boldsymbol{y}) \equiv
\left(
\begin{array}{c|c|c}
\boldsymbol{K} & \boldsymbol{N}^\top & \boldsymbol{T}^\top \\
\hline
-\rho \boldsymbol{D}_{\mathcal{A}_\nu} \boldsymbol{N} & \boldsymbol{D}_{\mathcal{I}_\nu} & \boldsymbol{0} \\
\hline
-\rho \boldsymbol{D}_{\mathcal{A}_t} \boldsymbol{T} & \mathcal{F}(\boldsymbol{D}_{\mathcal{I}_t^-} - \boldsymbol{D}_{\mathcal{I}_t^+}) & \boldsymbol{D}_{\mathcal{I}_t^+ \cup \mathcal{I}_t^-}
\end{array}
\right).
\tag{5}
$$

Note that the matrix $J(\boldsymbol{y})$ can be interpreted as a generalized Jacobi matrix namely, the differential of a slanting function related to the mapping G at the point $\boldsymbol{y}$, see [2]. We apply the *semi-smooth Newton method* for finding roots of (4).

<u>ALGORITHM *SSNM*</u>: Denote $\boldsymbol{F} \in \mathbb{R}^{n+2m}$, $\boldsymbol{F} \equiv (\boldsymbol{f}, 0, 0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, the right-hand side of (4). Set the tolerance $\varepsilon > 0$. Let $\boldsymbol{y}^{(0)} \in \mathbb{R}^{n+2m}$, $\rho > 0$, $k := 1$.

(i) Define the inactive/active sets related to $\boldsymbol{y}^{(k-1)}$. Assembly the relevant $J(\boldsymbol{y}^{(k-1)})$.

(ii) Compute $\boldsymbol{y}^{(k)}$ by solving the linear system

$$
J(\boldsymbol{y}^{(k-1)})\,\boldsymbol{y}^{(k)} = \boldsymbol{F}\,.
\tag{6}
$$

(iii) If $\|\boldsymbol{y}^{(k)} - \boldsymbol{y}^{(k-1)}\|/\|\boldsymbol{y}^{(k)}\| \le \varepsilon$, return $\boldsymbol{y} := \boldsymbol{y}^{(k)}$.

(iv) Set $k := k + 1$ and go to step (i).

In the case of convergence, we define

$$
\boldsymbol{y} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})
$$

as a numerical solution of problem (4). We usually set the tolerance $\varepsilon = 10^{-6}$, referring to the observation at the end of Section 1.

It is readily seen that if $\boldsymbol{y} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})$, $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, then

$$
(\boldsymbol{N}\boldsymbol{u})_i = 0,\ i \in \mathcal{A}_\nu, \quad (\boldsymbol{T}\boldsymbol{u})_i = 0,\ i \in \mathcal{A}_t,
\tag{7}
$$

$$
\lambda_{\nu,i} = 0,\ i \in \mathcal{I}_\nu, \quad \lambda_{t,i} + \mathcal{F}\lambda_{\nu,i} = 0,\ i \in \mathcal{I}_t^-, \quad \lambda_{t,i} - \mathcal{F}\lambda_{\nu,i} = 0,\ i \in \mathcal{I}_t^+.
\tag{8}
$$

As the active sets are complementary to the inactive sets, they define decoupling of contact nodes into two groups, i.e. the nodes with the Dirichlet conditions (7) and the nodes with the Neumann conditions (8).

Take another view: We may try to *guess* the inactive sets $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$ on the contact. Due to the dichotomy, it would imply the information concerning

$\mathcal{A} = \{\mathcal{A}_\nu; \mathcal{A}_t\}$. Hence, given $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$ on the contact, and given a load $\boldsymbol{f}$, find $(\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ such that

$$\left(\begin{array}{c|c|c} \boldsymbol{K} & \boldsymbol{N}^\top & \boldsymbol{T}^\top \\ \hline -\rho\boldsymbol{D}_{\mathcal{A}_\nu}\boldsymbol{N} & \boldsymbol{D}_{\mathcal{I}_\nu} & \boldsymbol{0} \\ \hline -\rho\boldsymbol{D}_{\mathcal{A}_t}\boldsymbol{T} & \mathcal{F}(\boldsymbol{D}_{\mathcal{I}_t^-} - \boldsymbol{D}_{\mathcal{I}_t^+}) & \boldsymbol{D}_{\mathcal{I}_t^+ \cup \mathcal{I}_t^-} \end{array} \right) \left(\begin{array}{c} \boldsymbol{u} \\ \hline \boldsymbol{\lambda}_\nu \\ \hline \boldsymbol{\lambda}_t \end{array} \right) = \left(\begin{array}{c} \boldsymbol{f} \\ \hline \boldsymbol{0} \\ \hline \boldsymbol{0} \end{array} \right) . \qquad (9)$$

System (9) can be interpreted as the discrete form of the Lamé equations (1) with the Dirichlet and Neumann boundary conditions (7) and (8), respectively. It motivates to define the linear operator

$$\boldsymbol{y} = DirNeu(\mathcal{I}, \boldsymbol{f}), \quad \boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) . \qquad (10)$$

Note that due to the clamping along Γ_u^1 and Γ_u^2, see Figure 1, the system (9) is uniquely solvable. The matrix $J(\boldsymbol{y})$ of this system is regular. This justifies, by the way, that iterations (6) are well defined.

Remark 3.1 Let $\boldsymbol{y}^{(0)} = DirNeu(\mathcal{I}, \boldsymbol{f})$. Then $\boldsymbol{y}^{(1)} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})$ and $\boldsymbol{y}^{(1)} = \boldsymbol{y}^{(0)}$ i.e., Algorithm *SSNM* converges in the first iteration. In other words, $\boldsymbol{y}^{(0)} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})$ is a fixed point of the iterations (6). Conversely, if $\boldsymbol{y}^{(0)} \in \mathbb{R}^{n+2m}$, $\boldsymbol{y}^{(0)} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})$, then defining $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$ to be the inactive sets of $\boldsymbol{y}^{(0)}$, we have $\boldsymbol{y}^{(0)} = DirNeu(\mathcal{I}, \boldsymbol{f})$. In that case, the solutions of the Dirichlet-Neumann problem (9) and the Coulomb friction problem (4) are identical.

Remark 3.2 In principle, we could find *all* roots $\boldsymbol{y}$ of (4) i.e., all fixed points $\boldsymbol{y}$ of the iterations (6). Given $\boldsymbol{f}$, make a trial choice of the inactive sets $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$ on the contact. Apply Remark 3.1: Let $\boldsymbol{y}^{(0)} = DirNeu(\mathcal{I}, \boldsymbol{f})$. The trial choice is successful, provided that $\boldsymbol{y}^{(0)} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f})$. The trouble is that we would have to check all $3 \sum_{j=0}^m \binom{m}{j} = 3 \cdot 2^m$ variants of the inactive sets $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$, which is not reasonable.

Remark 3.3 Let $\boldsymbol{y} = DirNeu(\mathcal{I}, \boldsymbol{f})$. The mapping G, see (4), is *not* differentiable at $\boldsymbol{y}$ provided that the active sets $\mathcal{A}_\nu$ and $\mathcal{A}_t$ have a special property: there exists a contact point $i \in \mathcal{M}$ such that

$$\text{either} \quad \lambda_{\nu,i} + \rho(\mathbf{N}\mathbf{u})_i = 0 \qquad (11)$$

$$\text{or} \quad \lambda_{t,i} + \rho(\mathbf{T}\mathbf{u})_i - \mathcal{F}\lambda_{\nu,i} = 0 \qquad (12)$$

$$\text{or} \quad \lambda_{t,i} + \rho(\mathbf{T}\mathbf{u})_i + \mathcal{F}\lambda_{\nu,i} = 0 . \qquad (13)$$

4. Continuation

Consider the Coulomb friction model (1)-(3), i.e. (4), assuming that $\boldsymbol{f} = \boldsymbol{f}(\alpha)$ depends on a scalar parameter α. We impose a continuous loading regime and seek for *continuous* response of the model. In particular, we consider a linear *loading path*

$$\boldsymbol{f}(\alpha) = (1 - \alpha)\boldsymbol{f}_1 + \alpha\boldsymbol{f}_2, \quad \alpha \in \mathbb{R}, \tag{14}$$

where $\boldsymbol{f}_1 \in \mathbb{R}^n$ and $\boldsymbol{f}_2 \in \mathbb{R}^n$ are given. The resulting *solution path* is a curve in $\mathbb{R} \times \mathbb{R}^{n+2m}$, see a qualitative sketch in Figure 2. It consists of *oriented linear branches*, connected by *transition points*.

Each oriented linear branch connecting transition points $(\alpha^{k-1}, \boldsymbol{y}^{k-1}) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ and $(\alpha^k, \boldsymbol{y}^k) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ is parameterized by α, and defined as

$$\alpha \longmapsto (\alpha, \boldsymbol{y}(\alpha)) \in \mathbb{R} \times \mathbb{R}^{n+2m}, \quad \boldsymbol{y}(\alpha) = DirNeu(\mathcal{I}, \boldsymbol{f}(\alpha)). \tag{15}$$

Note that the same branch (15) can have two different orientations. In particular,

- if $\alpha^{k-1} < \alpha^k$, we consider the positive orientation, i.e., $\alpha^{k-1} < \alpha < \alpha^k$, as α is increasing,

- if $\alpha^{k-1} > \alpha^k$, we consider the negative orientation, i.e., $\alpha^{k-1} > \alpha > \alpha^k$, as α is decreasing.

Let us emphasize that the inactive set $\mathcal{I}$ does not depend on the position of α in the above intervals. In Subsection 4.1, we give a predictor/corrector algorithm to follow such branch numerically. We can define the *orientation* of a particular branch by setting

$$s \equiv \frac{\alpha^k - \alpha^{k-1}}{|\alpha^k - \alpha^{k-1}|}.$$

Hence, orientation s attains the value $s = 1$ (positive orientation) and $s = -1$ (negative orientation). The mentioned predictor/corrector algorithm follows a branch with the same orientation s.

Oriented linear branch terminates in a transition point $(\alpha^k, \boldsymbol{y}^k) \in \mathbb{R} \times \mathbb{R}^{n+2m}$. It is related to a fixed point $\boldsymbol{y}^k = SSNM(\boldsymbol{y}^k, \boldsymbol{f}(\alpha^k))$. Due to Remark 3.1, $\boldsymbol{y}^k = DirNeu(\mathcal{I}, \boldsymbol{f}(\alpha^k))$, where $\mathcal{I} = \{\mathcal{I}_\nu; \mathcal{I}_t^+; \mathcal{I}_t^-\}$ are the inactive sets of $\boldsymbol{y}^k$. It can be shown that in a transition point $(\alpha^k, \boldsymbol{y}^k) \in \mathbb{R} \times \mathbb{R}^{n+2m}$, the mapping G, see (4), is *not* differentiable. We refer to Remark 3.3 for the analysis. Note that our objective is not to localize transition points exactly. In fact, due to rounding errors it is not possible. Instead, we develop computationally stable *branching* and *orientation* indicators which are formally related to each of the transition points, see Subsection 4.2.

4.1. Continuation of an oriented linear branch

Data of a linear branch: The orientation s and the fixed inactive set $\mathcal{I}$. The continuation algorithm is defined as a one-step recurrence

$$(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m} \to (\alpha_i, \boldsymbol{y}(\alpha_i)) \in \mathbb{R} \times \mathbb{R}^{n+2m}.$$
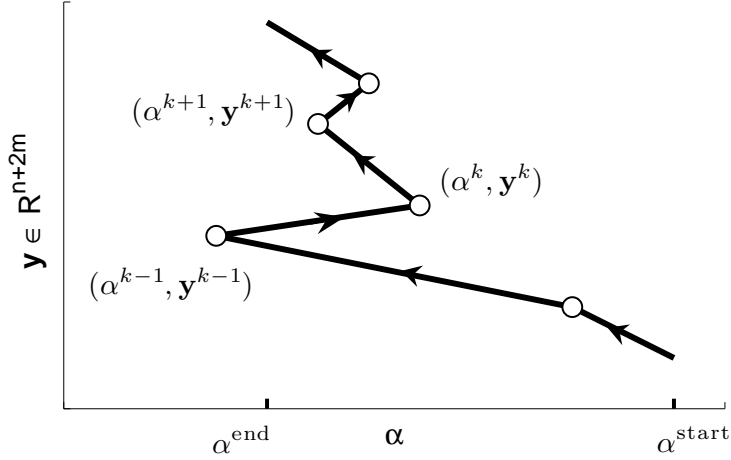
Figure 2: Solution path. For a fixed α, we may encounter up to five crossing points of the paths. They are related to five different solutions of equation (4) for the same right-hand side.

Parameters of the algorithm: The step-length h, in a range $0 < h_{\min} \leq h \leq h_{\max}$. The adaptive step-length strategy: Define c_s and c_p, $0 < c_s < 1 < c_p$, the *shortening* and the *prolongation* rates.

Let $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ be given. Consider the following

PREDICTOR-CORRECTOR ALGORITHM:

(i) *Predictor*: $\alpha^{new} = \alpha_{i-1} + sh$, $\quad \boldsymbol{y}^{(0)} = DirNeu(\mathcal{I}, \boldsymbol{f}(\alpha^{new}))$.

(ii) *Corrector*:

```
    if      y⁽¹⁾ = SSNM(y⁽⁰⁾, f(αⁿᵉʷ))   &   y⁽¹⁾ = y⁽⁰⁾
            return   αᵢ := αⁿᵉʷ, y(αᵢ) := y⁽¹⁾, i := i + 1, h := min(cₚh, hₘₐₓ)
    elseif  h < hₘᵢₙ
            return   continuation failed, the last computed point of the branch:
            (αᵢ₋₁, y(αᵢ₋₁)) with orientation s and the inactive set I
    else    h := max(cₛh, hₘᵢₙ),  go to step (i).
```

The algorithm returns either the new continuation point $(\alpha_i, \boldsymbol{y}(\alpha_i)) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ with the same orientation s and the inactive set $\mathcal{I}$, or fails - the case which will be discussed in Subsection 4.2.

Note that the above algorithm can be characterized as a *tangent continuation*, see [3], Algorithm 4.25. The step-size control is inspired by [4].

Figure 2: Solution path. For a fixed α, we may encounter up to five crossing points of the paths. They are related to five different solutions of equation (4) for the same right-hand side.

Parameters of the algorithm: The step-length h, in a range $0 < h_{\min} \leq h \leq h_{\max}$. The adaptive step-length strategy: Define c_s and c_p, $0 < c_s < 1 < c_p$, the *shortening* and the *prolongation* rates.

Let $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ be given. Consider the following

PREDICTOR-CORRECTOR ALGORITHM:

(i) *Predictor*: $\alpha^{new} = \alpha_{i-1} + sh$, $\quad \boldsymbol{y}^{(0)} = DirNeu(\mathcal{I}, \boldsymbol{f}(\alpha^{new}))$.

(ii) *Corrector*:

$$
\begin{aligned}
&\texttt{if} \quad \boldsymbol{y}^{(1)} = SSNM(\boldsymbol{y}^{(0)}, \boldsymbol{f}(\alpha^{new})) \quad \& \quad \boldsymbol{y}^{(1)} = \boldsymbol{y}^{(0)} \\
&\qquad \texttt{return} \quad \alpha_i := \alpha^{new}, \boldsymbol{y}(\alpha_i) := \boldsymbol{y}^{(1)}, i := i+1, h := \min(c_p h, h_{\max}) \\
&\texttt{elseif} \quad h < h_{\min} \\
&\qquad \texttt{return} \quad \text{continuation failed, the last computed point of the branch:} \\
&\qquad (\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \text{ with orientation } s \text{ and the inactive set } \mathcal{I} \\
&\texttt{else} \quad h := \max(c_s h, h_{\min}), \text{ go to step } (i).
\end{aligned}
$$

The algorithm returns either the new continuation point $(\alpha_i, \boldsymbol{y}(\alpha_i)) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ with the same orientation s and the inactive set $\mathcal{I}$, or fails - the case which will be discussed in Subsection 4.2.

Note that the above algorithm can be characterized as a *tangent continuation*, see [3], Algorithm 4.25. The step-size control is inspired by [4].
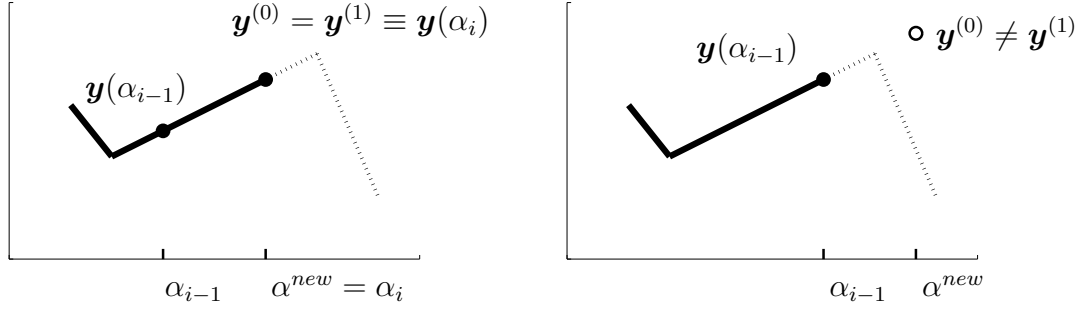
Figure 3: Oriented linear branch, predictor-corrector step. The corrector step is either accepted (on the left) or not accepted (on the right), and step-size h has to be shortened accordingly.

4.2. The branching and orientation indicators

Let $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ be the last point of a linear branch with an orientation s and inactive set $\mathcal{I}$, see the failure of path following the linear branch in Subsection 4.1. Define a trial point $(\alpha^{\text{fail}}, \boldsymbol{y}^{\text{fail}}) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ setting

$$\alpha^{\text{fail}} = \alpha_{i-1} + s h_{\text{fail}}, \quad \boldsymbol{y}^{\text{fail}} = DirNeu(\mathcal{I}, \boldsymbol{f}(\alpha^{\text{fail}})), \tag{16}$$

where h_{fail} is the step-length related to the failure of continuation. Note that $\boldsymbol{y}^{\text{fail}} \neq SSNM(\boldsymbol{y}^{\text{fail}}, \boldsymbol{f}(\alpha^{\text{fail}}))$. Figure 4, the upper panel, suggests that $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1}))$ and $(\alpha^{\text{fail}}, \boldsymbol{y}^{\text{fail}})$ are close to a transition point. We may envisage two qualitatively different cases of the transition.

According to the generic scenario, we should indicate a change of $\mathcal{I}$: Let $\boldsymbol{u}$, $\boldsymbol{\lambda}_\nu$ and $\boldsymbol{\lambda}_t$ denote the solution components $\boldsymbol{y}(\alpha_{i-1}) = (\boldsymbol{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$. Let

$$M = \min \left\{ |\boldsymbol{\lambda}_\nu + \rho \boldsymbol{N} \boldsymbol{u}|, \ |\boldsymbol{\lambda}_t + \rho \boldsymbol{T} \boldsymbol{u} - \mathcal{F} \boldsymbol{\lambda}_\nu|, \ |\boldsymbol{\lambda}_t + \rho \boldsymbol{T} \boldsymbol{u} + \mathcal{F} \boldsymbol{\lambda}_\nu| \right\}. \tag{17}$$

The idea is that the minimizer of the above expression should indicate a transition point. We expect that just one component of the minimizer is significant. The transition is related to a transition between *inactive* and *active* sets. In this respect, the minimizer is interpreted as follows:

$$\left. \begin{array}{llll} \text{If} & M = |\boldsymbol{\lambda}_\nu + \rho \boldsymbol{N} \boldsymbol{u}|_i, & \text{then} & \mathcal{A}_\nu \overset{i}{\longleftrightarrow} \mathcal{I}_\nu \\ \text{else if} & M = |\boldsymbol{\lambda}_t + \rho \boldsymbol{T} \boldsymbol{u} - \mathcal{F} \boldsymbol{\lambda}_\nu|_i, & \text{then} & \mathcal{A}_t \overset{i}{\longleftrightarrow} \mathcal{I}_t^+ \\ \text{else if} & M = |\boldsymbol{\lambda}_t + \rho \boldsymbol{T} \boldsymbol{u} + \mathcal{F} \boldsymbol{\lambda}_\nu|_i, & \text{then} & \mathcal{A}_t \overset{i}{\longleftrightarrow} \mathcal{I}_t^- \end{array} \right\} \quad \mathcal{I}_{new} := \mathcal{I}. \tag{18}$$

The symbol "$\overset{i}{\longleftrightarrow}$" indicates a particular transition of the index i between the active and the inactive set. The procedure above results in an update of $\mathcal{I}$ denoted as $\mathcal{I}_{new}$.
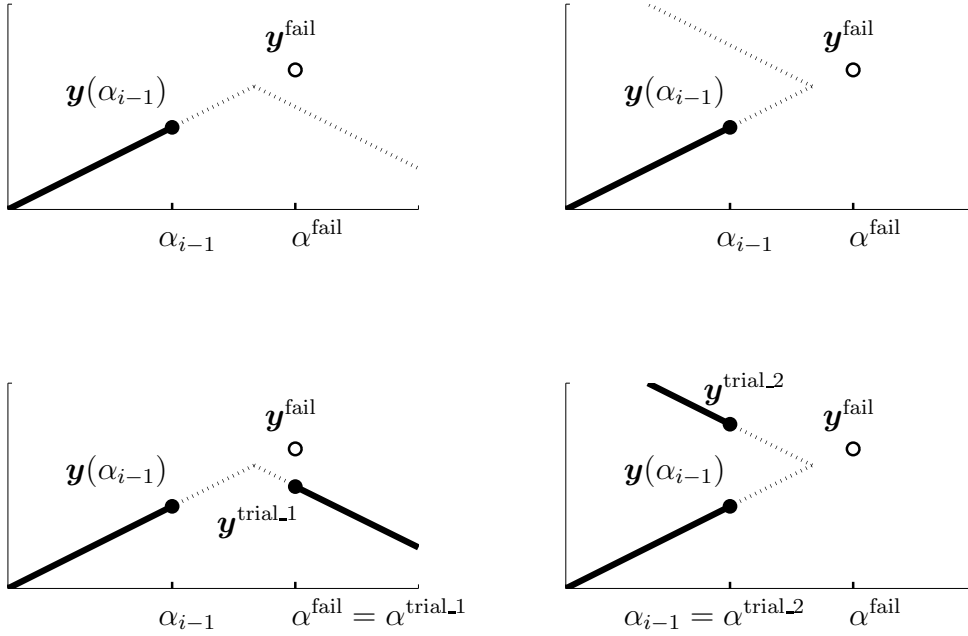
We propose the following

Figure 4: Let $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ be the last point on a linear branch, continuation failure indicated on $(\alpha^{\mathrm{fail}}, \boldsymbol{y}(\alpha^{\mathrm{fail}})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$. The upper panel, qualitative scenario envisaged: a) transversal transition on the left, b) fold (turning point) transition on the right. The lower panel: Branching due to the algorithm.

<u>Branching Algorithm</u>

Let $(\alpha_{i-1}, \boldsymbol{y}(\alpha_{i-1})) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ be the last point of a linear branch with an orientation s and inactive set $\mathcal{I}$. Update $\mathcal{I}_{new}$ via (18).

Define $\alpha^{\mathrm{trial\_1}} = \alpha_{i-1} + sh_{\mathrm{fail}}$ and $\boldsymbol{y}^{\mathrm{trial\_1}} = DirNeu(\mathcal{I}_{new}, \boldsymbol{f}(\alpha^{\mathrm{trial\_1}}))$.

`If`
$\boldsymbol{y}^{\mathrm{trial\_1}} = SSNM(\boldsymbol{y}^{\mathrm{trial\_1}}, \boldsymbol{f}(\alpha^{\mathrm{trial\_1}}))$, set $(\alpha^{\mathrm{trial\_1}}, \boldsymbol{y}^{\mathrm{trial\_1}}) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ to be the first point on a linear branch with the orientation $s := s$ and the inactive set $\mathcal{I} := \mathcal{I}_{new}$.

Comment: *transversal transition.*

`else`
Define $\alpha^{\mathrm{trial\_2}} = \alpha_{i-1}$ and $\boldsymbol{y}^{\mathrm{trial\_2}} = DirNeu(\mathcal{I}_{new}, \boldsymbol{f}(\alpha^{\mathrm{trial\_2}}))$.
Set $(\alpha^{\mathrm{trial\_2}}, \boldsymbol{y}^{\mathrm{trial\_2}}) \in \mathbb{R} \times \mathbb{R}^{n+2m}$ to be the first point of a linear branch with orientation $s := -s$ and inactive set $\mathcal{I} := \mathcal{I}_{new}$.
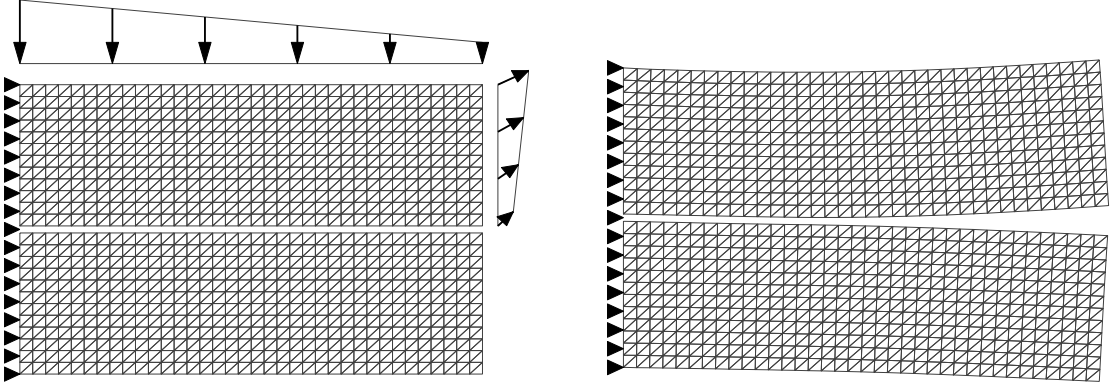
Comment: *fold, turning point.*

113

Figure 5: Contact of two elastic bodies Ω^1 (the upper body) and Ω^2, along the contact boundary Γ_c. The loading is due to the surface traction. On the right: Resulting deformation.

Set $i := i + 1$, and apply continuation of the oriented linear branch with orientation s and the inactive set $\mathcal{I}$.

The idea of the algorithm is indicated in the lower panel of Figure 4. The algorithm works provided that $h_{\min}$ is sufficiently small.

The ambition of the present paper is not to justify the branching scenario theoretically. Note only, that the transversal transition may be described using a proper version of the Implicit Function Theorem, see e.g. [15, 5] and [8] in the context of Coulomb friction. In case of the fold transition, we cannot quote (to our knowledge) a relevant analytical result immediately.

5. Numerical experiments

We consider a particular geometry, see Figure 5.

The actual computations are depicted in Figure 6. If $\mathcal{F}$ is sufficiently small then the solution path should contain transversal transition points only, see e.g. [8]. It pertains to Figure 6, upper left. For larger friction coefficients (e.g. $\mathcal{F} = 0.6$ and $\mathcal{F} = 30$), the path-following algorithm reveals non-unique solutions of the problem, see Figure 6, upper right and lower-left including the corresponding zoom. In particular, we can have up to three ($\mathcal{F} = 0.6$) and five ($\mathcal{F} = 30$) solutions for a fixed parameter α.
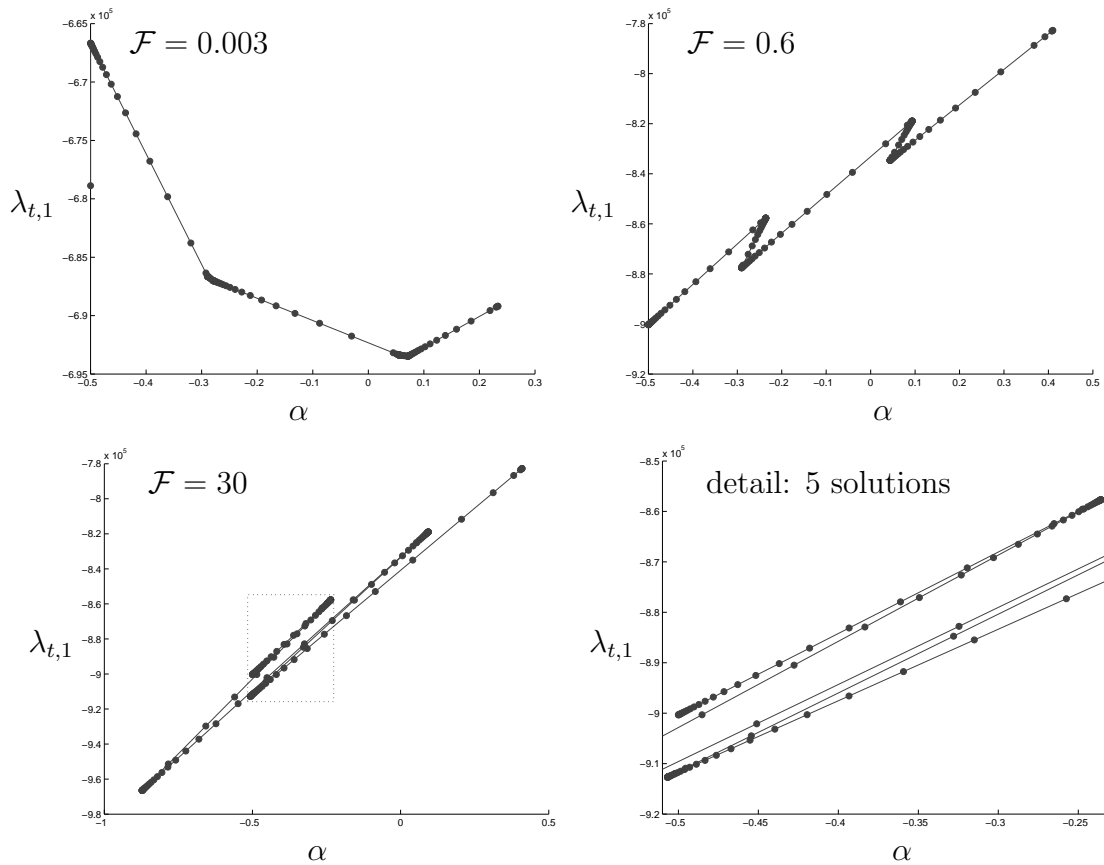
Acknowledgements

Figure 6: Discretization: $n = 1320$, $m = 30$. The stepsize control: $10^{-5} \leq h \leq 5$, $c_p = 1.3$, $c_s = 0.5$. Plots: Parameter α vs. the solution component $\lambda_{t,1}$, for selected friction coefficients $\mathcal{F}$.

References

[1] Allgower, E. L. and Georg, K.: *Numerical path following.* Handbook of Numerical Analysis, vol. V, Elsevier Science, New York, 1997.

[2] Chen, X., Nashed, Z., and Qi, L.: Smoothing methods and semismooth methods for nondifferential operator equations. SIAM J. Numer. Anal. **5** (2000), 1200–1216.

[3] Deuflhart, P. and Hohmann, A.: *Numerical analysis in modern scientific computing.* Texts in Applied Mathematics, Springer Verlag, New York, 2003.

[4] Dhooge, A., Govaerts, W., and Kuzetsov, Y. A.: MATCONT: A Matlab package for numerical bifurcation analysis of ODEs. ACM Trans. Math. Software **31** (2003), 141–164.

[5] Dontchev, A. L. and Rockafellar, R. T.: Robinson's implicit function theorem and its extensions. Math. Program. **117** (2009), 129–147.

[6] Eck, C. and Jarušek, J.: Existence results for the static contact problems with Coulomb friction. Math. Models Methods Appl. Sci. **8** (1997), 445–468.

[7] Facchinei, F. and Pang, J.: *Finite-dimensional variational inequalities and complementarity problems*. Springer Series in Operations Research xxxiii, Springer Verlag, New York, 2003.

[8] Haslinger, J., Janovský, V., and Ligurský, T.: Qualitative analysis of solutions to discrete static contact problems with Coulomb friction. Comput. Meth. Appl. Mech. Engrg. **205–208** (2012), 149–161.

[9] Hild, P. and Renard, Y.: Local uniqueness and continuation of solutions for the discrete Coulomb friction problem in elastostatics. Quart. Appl. Math. **63** (2005), 553–573.

[10] Ito, K. and Kunisch, K.: Semi-smooth Newton methods for the Signorini problem. Appl. Math. **53** (2009), 455–468.

[11] Janovský, V.: Catastrophic features of Coulomb friction model. In: J. R. Whiteman (Ed.), *The Mathematics of Finite Elements and Applications IV*, pp. 259–264. Academic Press, New York, 1982.

[12] Janovský, V. and Ligurský, T.: Computing non unique solutions of the Coulomb friction problem. Math. Comput. Simulation **82** (2012), 2047–2061.

[13] Ligurský, T.: Theoretical analysis of discrete contact problems with Coulomb friction. Appl. Math. **57** (2012), 263–295.

[14] Nečas, J., Jarušek, J., and Haslinger, J.: On the solution of variational inequality to the Signorini problem with small friction. Bolletino U.M.I. **5** (1980), 796–811.

[15] Robinson, S.: An implicit-function theorem for a class of nonsmooth functions. Math. Oper. Res. **16** (1991), 292–309.

[16] Scholtes, S.: *Introduction to piecewise differentiable equations*. SpringerBriefs in Optimalization, Springer, Berlin, 2012.