Bifurcations in Piecewise-Smooth Steady-State Problems

Tomas Ligursky, Yves Renard

Computational Mechanics manuscript No.
(will be inserted by the editor)

Bifurcations in Piecewise-Smooth Steady-State Problems

Abstract Study and Application to Plane Contact Problems with Friction

T. Ligurský · Y. Renard

Received: October 24, 2014

Abstract The paper presents a local study of bifurcations in a class of piecewise-smooth steady-state problems for which the regions of smooth behaviour permit analytical expressions. A system of piecewise-linear equations capturing the essential features of the branching scenarios around points of non-smoothness is derived under the assumptions that (i) the points lie in the intersection of the boundaries of the regions where the gradients of the respective smooth selections have the full rank, (ii) there is no solution branch whose tangential direction is tangent to the boundary of any of the regions. The simplest cases of this system are studied in detail and the most probable branching scenarios are described. A criterion for detecting bifurcation points is proposed and a procedure for its realisation in the course of numerical continuation of solution curves of large problems is designed. Application of the general frame to discretised plane contact problems with Coulomb friction is explained. Simple as well as more realistic model examples of bifurcations are shown.

Keywords Piecewise-smooth · Steady-state problem · Local bifurcation · Bifurcation criterion · Contact problem · Coulomb friction

T. Ligurský
Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic
Tel.: +420-585634606
E-mail: tomas.ligursky@upol.cz

Y. Renard
Université de Lyon, CNRS,
INSA-Lyon, ICJ UMR5208, F 69621, Villeurbanne, France.
E-mail: Yves.Renard@insa-lyon.fr

1 Introduction

The steady-state bifurcation problem:

$$\left.\begin{array}{l}\textit{Find } \boldsymbol{y} \in U \textit{ such that}\\ \boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{0},\end{array}\right\} \qquad (\mathscr{P})$$

where $U \subset \mathbb{R}^{N+1}$ and $\boldsymbol{H}\colon U \to \mathbb{R}^N$, has been the subject of large number of studies in the last decades (see, e.g., [4, Section 24] and the references therein). If $\boldsymbol{H}$ is smooth, say continuously differentiable, this problem is quite well understood from the theoretical point of view and a great variety of methods has been constructed for its numerical treatment.

On the other hand, there are many equilibrium problems in economics and diverse engineering fields, whose models lead naturally to a system of non-smooth equations [3, 27, 14]. For instance, let us mention discretised frictionless and frictional contact problems in solid mechanics, which are of our specific interest. In general, important classes of variational inequalities, complementarity problems and constrained optimisation problems can be reformulated in this way.

Nevertheless, the question of bifurcations of solutions of such problems when they depend on a parameter is very much open to our knowledge: The local existence and first-order approximation of branches of solutions of variational inequalities were studied and methods of numerical continuation of the branches were proposed in [10, 9, 23]. But the subject of branching was not touched at all there. The papers [25, 29, 22] deal with the analysis of bifurcations in constrained optimisation problems, their numerical detection as well as branch switching. However, difficulties related to non-smoothness are circumvented to great extent by considering a set of points that are more general than stationary ones. Bifurcations of static equilibrium curves of frictionless contact problems were studied in [7, 28, 20], where the tangential directions of curves emanating from points of non-smoothness were determined by a certain mixed linear complementarity problem and a method based on resolution of this problem was proposed for branch switching during numerical continuation.

In our recent paper [21], we analysed local continuation of solution curves of Problem ($\mathscr{P}$) with $\boldsymbol{H}$ piecewise C^1 (PC^1) and we developed a method of numerical continuation for this case. We established also some particular results for plane contact problems with Coulomb friction, whose formulation as a non-smooth system fits perfectly the PC^1-setting. Techniques of numerical continuation for frictional plane contact problems can be found also in [19, 17, 16]. But no special care was taken of bifurcation points in these papers.

Besides, there is a vast amount of existing literature on bifurcations in piecewise-smooth *dynamical* systems (see [6] and the references therein). Since much more phenomena can be observed in dynamical systems than in the steady-state ones, the literature is restricted almost entirely to bifurcations occurring on boundaries between two regions of smooth behaviour of the function involved. This situation is not of much interest in the steady-state case as we shall show later on.

The present paper deals with the case of Problem ($\mathscr{P}$) where $\boldsymbol{H}$ belongs to a class of PC^1-functions with analytical expressions for the regions of smooth behaviour (see Assumption 1 for the precise definition). It focuses on branching occurring in solutions lying in intersections of the boundaries of two or more regions. By exploiting the structure of the set of points of non-smoothness, we complete our local description of the solution set around such a boundary solution from [21] under certain non-degeneracy assumptions. It is worth mentioning that despite the conditions imposed in our definition, the considered subclass of PC^1-functions remains still quite general and suitable for the formulation of plane contact problems with Coulomb friction.

The outline of our study is the following: In Section 2, bifurcation points of ($\mathscr{P}$) are analysed in general and a system of piecewise-linear equations capturing the essential features of the branching scenarios around certain types of boundary solutions, the so-called border-collision solutions, is derived. Its simplest cases are then studied in detail, the most probable branching scenarios are shown and a bifurcation criterion based on our observations is proposed. The aim of Section 3 is to present a procedure for realisation of this criterion in the course of numerical continuation. Application of the general frame to plane contact problems with friction is demonstrated in Section 4. Finally, model examples of bifurcations are shown in Section 5.

The following notation is used throughout the paper: The interior of a set A is denoted by $\mathring{A}$ or $\operatorname{int} A$, the closure by $\overline{A}$, the boundary by ∂A, the exterior by $\operatorname{ext} A$ and the orthogonal complement by $A^\perp$. The gradients of a real-valued function f and a vector-valued function $\boldsymbol{f}$ at a point $\bar{\boldsymbol{x}}$ are written as $\nabla f(\bar{\boldsymbol{x}})$ and $\boldsymbol{\nabla f}(\bar{\boldsymbol{x}})$, respectively. If $\boldsymbol{f}$ is a function of two variables $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{\nabla_x f}(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$ stands for the partial gradient of $\boldsymbol{f}$ with respect to $\boldsymbol{x}$ at $(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$. We use systematically the convention that x_i and f_i are the ith component of a vector $\boldsymbol{x}$ and the ith component function of a vector-valued function $\boldsymbol{f}$, respectively.

Furthermore, let us recall essentials from theory of PC^1-functions from [27]:

Definition 1 A function $\boldsymbol{H}\colon U \to \mathbb{R}^N$, $U \subset \mathbb{R}^M$, is PC^1 if it is continuous and for every $\bar{\boldsymbol{y}} \in U$, there exist an open neighbourhood $O \subset U$ of $\bar{\boldsymbol{y}}$ and a finite family of C^1-functions $\boldsymbol{H}^{(i)}\colon O \to \mathbb{R}^N$, $i \in \mathscr{I}(\bar{\boldsymbol{y}})$, such that

$$\forall \boldsymbol{y} \in O\colon\ \boldsymbol{H}(\boldsymbol{y}) \in \{\boldsymbol{H}^{(i)}(\boldsymbol{y});\, i \in \mathscr{I}(\bar{\boldsymbol{y}})\}.$$

The functions $\boldsymbol{H}^{(i)}$ are termed *selections* of $\boldsymbol{H}$ at $\bar{\boldsymbol{y}}$.

One can show that every PC^1-function is B-differentiable, that is, it is directionally differentiable and the directional derivative of $\boldsymbol{H}$ at $\boldsymbol{y}$ in the direction $\boldsymbol{z}$ (denoted by $\boldsymbol{H}'(\boldsymbol{y};\boldsymbol{z})$) satisfies:

$$\lim_{\boldsymbol{z}\to\boldsymbol{0}} \frac{\|\boldsymbol{H}(\boldsymbol{y}+\boldsymbol{z}) - \boldsymbol{H}(\boldsymbol{y}) - \boldsymbol{H}'(\boldsymbol{y};\boldsymbol{z})\|}{\|\boldsymbol{z}\|} = 0.$$

A special case of PC^1-functions are *piecewise-linear* functions. These are continuous functions whose selections are linear, that is, of the form $\boldsymbol{y} \mapsto \boldsymbol{A}^{(i)}\boldsymbol{y}$ for some matrices $\boldsymbol{A}^{(i)}$. It holds that the directional derivative $\boldsymbol{H}'(\boldsymbol{y};.)$ of a PC^1-function $\boldsymbol{H}$ is a piecewise-linear function.

2 Theoretical Analysis

Our study of bifurcations of Problem ($\mathscr{P}$) is restricted to the functions $\boldsymbol{H}$ specified by (see Fig. 1 for illustration):

Assumption 1 *Let $\boldsymbol{H}\colon U \to \mathbb{R}^N$ be a PC^1-function such that every $\bar{\boldsymbol{y}} \in U$ meets one of the following two conditions:*
(i) *$\boldsymbol{H}$ is a C^1-function in an open neighbourhood O of $\bar{\boldsymbol{y}}$;*
(ii) *there exist an open neighbourhood O of $\bar{\boldsymbol{y}}$ and finite families of C^1-selections $\boldsymbol{H}^{(i)}$ and regions $D^{(i)}$, $i \in \mathscr{I}(\bar{\boldsymbol{y}})$, satisfying*

$$\boldsymbol{H} = \boldsymbol{H}^{(i)} \text{ in } D^{(i)}, \quad D^{(i)} = \{\boldsymbol{y} \in O;\, \boldsymbol{G}^{(i)}(\boldsymbol{y}) \le \boldsymbol{0}\} \tag{1}$$

for some C^1-functions $\boldsymbol{G}^{(i)}\colon O \to \mathbb{R}^{M_i}$. Moreover,

$$\boldsymbol{G}^{(i)}(\bar{\boldsymbol{y}}) = \boldsymbol{0}, \quad \forall i \in \mathscr{I}(\bar{\boldsymbol{y}}), \tag{2}$$

$$\nabla G_j^{(i)}(\bar{\boldsymbol{y}}) \ne \boldsymbol{0}, \quad \forall i \in \mathscr{I}(\bar{\boldsymbol{y}}),\ j = 1,\dots,M_i, \tag{3}$$

$$\bigcup_{i\in\mathscr{I}(\bar{\boldsymbol{y}})} D^{(i)} = O, \tag{4}$$

$$\mathring{D}^{(i)} \cap \mathring{D}^{(j)} = \emptyset, \quad \forall i,j \in \mathscr{I}(\bar{\boldsymbol{y}}),\ i \ne j. \tag{5}$$

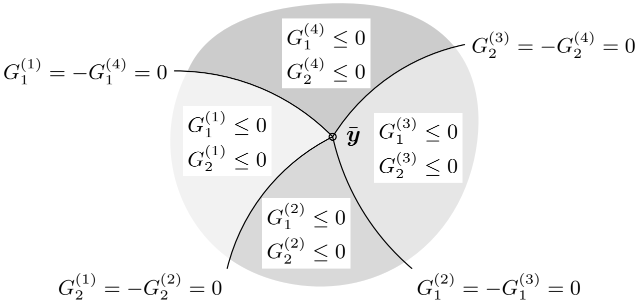


Fig. 1 Example of regions of smoothness of $\boldsymbol{H}$ from Assumption 1(ii).

Let us note that the function class determined by this assumption is very close to the ones termed piecewise smooth in [2, Appendix I] and [3]. In particular, the expression (1) is natural for PC^1-functions whose component functions are of max-min type or involve projections onto intervals (see [3] and Section 4 for examples).

To start with our analysis, we modify slightly the definition of a bifurcation point for smooth functions from [4].

Definition 2 Let J be an open interval containing $\bar{s}$, and $\boldsymbol{c}\colon J \to \mathbb{R}^{N+1}$ be a curve such that $\boldsymbol{H}(\boldsymbol{c}(s)) = \boldsymbol{0}$ for any $s \in J$. The point $\boldsymbol{c}(\bar{s})$ is called a *bifurcation point* of the equation $\boldsymbol{H} = \boldsymbol{0}$ if there exists an $\varepsilon > 0$ such that every neighbourhood of $\boldsymbol{c}(\bar{s})$ contains zero points of $\boldsymbol{H}$ that are not on $\boldsymbol{c}(\bar{s}-\varepsilon, \bar{s}+\varepsilon)$.

Let $\bar{\boldsymbol{y}}$ be a known solution of $(\mathscr{P})$. The classical implicit-function theorem guarantees that if there is only one selection of $\boldsymbol{H}$ at $\bar{\boldsymbol{y}}$, that is, $\boldsymbol{H}$ is smooth there, and its gradient has the full row rank, then there exists locally a unique (smooth) solution curve. Hence, $\bar{\boldsymbol{y}}$ may be a bifurcation point only in one of the following three cases:

(i) The gradient of the only selection is rank-deficient at $\bar{\boldsymbol{y}}$.
(ii) There are two or more selections at $\bar{\boldsymbol{y}}$ and all of them have a full-row rank gradient at $\bar{\boldsymbol{y}}$.
(iii) There are two or more selections at $\bar{\boldsymbol{y}}$ and at least one of them has a rank-deficient gradient at $\bar{\boldsymbol{y}}$.

The first case leads to well-established theory of smooth bifurcations whereas in the second case, the solution set of $(\mathscr{P})$ is composed of parts of unique solution curves of $\boldsymbol{H}^{(i)} = \boldsymbol{0}$ for individual selections $\boldsymbol{H}^{(i)}$ around $\bar{\boldsymbol{y}}$. The third case can be viewed as a combination of the first two ones, where some selections may contribute to the whole solution set with more than one solution curve, and this seems to be the most rare. This motivates us to focus on the second case, that is, *purely non-smooth* bifurcations. Henceforth, we shall consider a fixed solution $\bar{\boldsymbol{y}}$ lying on the boundary of two or more regions $D^{(i)}$ from Assumption 1(ii), a so-called *boundary solution*, and we shall assume the following:

Assumption 2 *The gradient $\nabla \boldsymbol{H}^{(i)}(\bar{\boldsymbol{y}})$ of any selection $\boldsymbol{H}^{(i)}$, $i \in \mathscr{I}(\bar{\boldsymbol{y}})$, has the full rank.*

2.1 Border-Collision Normal Form

Following [6, Subsection 3.1.3] and imposing another non-degeneracy condition, which will be specified later on, we shall derive a simplified system that captures the essential features of the branching scenarios around the boundary solution $\bar{\boldsymbol{y}}$. The system will consist of piecewise-linear equations.

Firstly, we introduce new co-ordinates $\tilde{\boldsymbol{y}} := \boldsymbol{y} - \bar{\boldsymbol{y}}$ and pass to

$$\tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) := \boldsymbol{H}(\tilde{\boldsymbol{y}} + \bar{\boldsymbol{y}}), \quad \tilde{\boldsymbol{H}}^{(i)}(\tilde{\boldsymbol{y}}) := \boldsymbol{H}^{(i)}(\tilde{\boldsymbol{y}} + \bar{\boldsymbol{y}}),$$
$$\tilde{\boldsymbol{G}}^{(i)}(\tilde{\boldsymbol{y}}) := \boldsymbol{G}^{(i)}(\tilde{\boldsymbol{y}} + \bar{\boldsymbol{y}}), \quad \tilde{O} := O - \{\bar{\boldsymbol{y}}\},$$
$$\tilde{D}^{(i)} := \{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{G}}^{(i)}(\tilde{\boldsymbol{y}}) \le \boldsymbol{0}\}$$

so that the boundary solution $\bar{\boldsymbol{y}}$ is translated to $\boldsymbol{0}$.

Since every PC^1-function is B-differentiable, we can expand $\tilde{\boldsymbol{H}}$ about $\boldsymbol{0}$ as

$$\tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) = \tilde{\boldsymbol{H}}(\boldsymbol{0}) + \tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}}) + \boldsymbol{o}(\tilde{\boldsymbol{y}}).$$

Inserting $\tilde{\boldsymbol{H}}(\boldsymbol{0}) = \boldsymbol{0}$ and neglecting the term $\boldsymbol{o}(\tilde{\boldsymbol{y}})$, we obtain the following simplification of $(\mathscr{P})$:

$$\tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}}) = \boldsymbol{0}. \tag{6}$$

Next, we shall give an explicit expression for $\tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}})$. For this purpose, we set

$$\boldsymbol{A}^{(i)} := \nabla \tilde{\boldsymbol{H}}^{(i)}(\boldsymbol{0}) = \nabla \boldsymbol{H}^{(i)}(\bar{\boldsymbol{y}}), \tag{7}$$
$$\boldsymbol{B}^{(i)} := \nabla \tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0}) = \nabla \boldsymbol{G}^{(i)}(\bar{\boldsymbol{y}}), \tag{8}$$
$$C^{(i)} := \{\tilde{\boldsymbol{y}} \in \mathbb{R}^{N+1};\ \boldsymbol{B}^{(i)}\tilde{\boldsymbol{y}} \le \boldsymbol{0}\}, \tag{9}$$
$$\mathscr{I}' := \{i \in \mathscr{I}(\bar{\boldsymbol{y}});\ \mathring{C}^{(i)} \ne \emptyset\}. \tag{10}$$

Theorem 1 *Under Assumption 1,*

$$\bigcup_{i \in \mathscr{I}'} C^{(i)} = \mathbb{R}^{N+1}, \tag{11}$$
$$\mathring{C}^{(i)} = \{\tilde{\boldsymbol{y}} \in \mathbb{R}^{N+1};\ \boldsymbol{B}^{(i)}\tilde{\boldsymbol{y}} < \boldsymbol{0}\}, \quad \forall i \in \mathscr{I}', \tag{12}$$
$$\mathring{C}^{(i)} \cap C^{(j)} = \emptyset, \quad \forall i,j \in \mathscr{I}',\ i \ne j, \tag{13}$$
$$\tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}}) = \boldsymbol{A}^{(i)}\tilde{\boldsymbol{y}}, \quad \forall \tilde{\boldsymbol{y}} \in C^{(i)},\ \forall i \in \mathscr{I}'. \tag{14}$$

Proof First, we shall show that

$$\bigcup_{i \in \mathscr{I}(\bar{\boldsymbol{y}})} C^{(i)} = \mathbb{R}^{N+1}. \tag{15}$$

Let $\tilde{\boldsymbol{y}} \in \mathbb{R}^{N+1}$ be arbitrarily chosen. Since $\tilde{O}$ is a neighbourhood of $\boldsymbol{0}$, which is covered by $\{\tilde{D}^{(i)}\}_{i \in \mathscr{I}(\bar{\boldsymbol{y}})}$ by (4), and $\mathscr{I}(\bar{\boldsymbol{y}})$ is finite, there exists $i_0 \in \mathscr{I}(\bar{\boldsymbol{y}})$ and $\{r_n\} \subset \mathbb{R}$, $r_n \to 0+$, such that $r_n\tilde{\boldsymbol{y}} \in \tilde{D}^{(i_0)}$ for any n, that is, $\tilde{\boldsymbol{G}}^{(i_0)}(r_n\tilde{\boldsymbol{y}}) \le 0$. From here and (2),

$$\boldsymbol{B}^{(i_0)}\tilde{\boldsymbol{y}} = \nabla \tilde{\boldsymbol{G}}^{(i_0)}(\boldsymbol{0})\tilde{\boldsymbol{y}} = \lim_{n\to\infty} \frac{\tilde{\boldsymbol{G}}^{(i_0)}(r_n\tilde{\boldsymbol{y}}) - \tilde{\boldsymbol{G}}^{(i_0)}(\boldsymbol{0})}{r_n} \le 0,$$

that is, $\tilde{\boldsymbol{y}} \in C^{(i_0)}$. This yields (15).

Second, suppose that $\mathring{C}^{(i_0)} = \emptyset$ and $\tilde{\boldsymbol{y}} \in C^{(i_0)}$. Then due to (15),

$$\tilde{\boldsymbol{y}} \in \overline{\bigcup_{i \in \mathscr{I}(\bar{\boldsymbol{y}}) \setminus \{i_0\}} C^{(i)}} = \bigcup_{i \in \mathscr{I}(\bar{\boldsymbol{y}}) \setminus \{i_0\}} \overline{C^{(i)}} = \bigcup_{i \in \mathscr{I}(\bar{\boldsymbol{y}}) \setminus \{i_0\}} C^{(i)}$$

and (11) follows by induction.

To prove (12), it suffices to verify

$$\mathring{C}^{(i)} \subset \{\tilde{\boldsymbol{y}} \in \mathbb{R}^{N+1};\ \boldsymbol{B}^{(i)}\tilde{\boldsymbol{y}} < \boldsymbol{0}\}, \quad i \in \mathscr{I}'.$$

Suppose for contradiction that there is $\tilde{\boldsymbol{y}} \in \mathring{C}^{(i)}$ and an index $j \in \{1, \ldots, M_i\}$ with $\boldsymbol{B}_j^{(i)}\tilde{\boldsymbol{y}} = 0$, where $\boldsymbol{B}_j^{(i)}$ stands for the jth row of $\boldsymbol{B}^{(i)}$. Then one can find a neighbourhood $\tilde{V}$ of $\tilde{\boldsymbol{y}}$ such that for any $\tilde{\boldsymbol{z}} \in \tilde{V}$,

$$\boldsymbol{B}_j^{(i)}\tilde{\boldsymbol{z}} \le 0,$$

$$\boldsymbol{B}_j^{(i)}(\tilde{\boldsymbol{z}} - \tilde{\boldsymbol{y}}) \le 0.$$

This implies

$$\boldsymbol{B}_j^{(i)}\boldsymbol{z} = 0, \quad \forall \boldsymbol{z} \in \mathbb{R}^{N+1},$$

which contradicts to (3), and (12) is valid.

Next, take $\tilde{\boldsymbol{y}} \in \mathring{C}^{(i)}$, $i \in \mathscr{I}'$, and $r > 0$ sufficiently small. From the mean-value theorem,

$$\begin{aligned}\big\|\tilde{\boldsymbol{G}}^{(i)}(r\tilde{\boldsymbol{y}}) - \tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0}) - \nabla\tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0})r\tilde{\boldsymbol{y}}\big\| &\\ \le \sup_{t\in[0,1]} \big\|\big(\nabla\tilde{\boldsymbol{G}}^{(i)}(tr\tilde{\boldsymbol{y}}) - \nabla\tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0})\big)r\tilde{\boldsymbol{y}}\big\| &\\ \le r \sup_{t\in[0,1]} \big\|\nabla\tilde{\boldsymbol{G}}^{(i)}(tr\tilde{\boldsymbol{y}}) - \nabla\tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0})\big\| \|\tilde{\boldsymbol{y}}\|, &\end{aligned}$$

and the continuous differentiability of $\tilde{\boldsymbol{G}}^{(i)}$ together with (2) ensures that

$$\begin{aligned}\tilde{\boldsymbol{G}}^{(i)}(r\tilde{\boldsymbol{y}}) &= \tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0}) + \nabla\tilde{\boldsymbol{G}}^{(i)}(\boldsymbol{0})r\tilde{\boldsymbol{y}} + \boldsymbol{o}(r)\\ &= r\big(\boldsymbol{B}^{(i)}\tilde{\boldsymbol{y}} + \boldsymbol{o}(1)\big).\end{aligned} \tag{16}$$

This and (12) yield that for any $r > 0$ small enough, $\tilde{\boldsymbol{G}}^{(i)}(r\tilde{\boldsymbol{y}}) < 0$, that is, $r\tilde{\boldsymbol{y}} \in \tilde{D}^{(i)}$. Hence,

$$\begin{aligned}\tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}}) &= \lim_{r\to 0+} \frac{\tilde{\boldsymbol{H}}(r\tilde{\boldsymbol{y}}) - \tilde{\boldsymbol{H}}(\boldsymbol{0})}{r}\\ &= \lim_{r\to 0+} \frac{\tilde{\boldsymbol{H}}^{(i)}(r\tilde{\boldsymbol{y}}) - \tilde{\boldsymbol{H}}^{(i)}(\boldsymbol{0})}{r} = \nabla\tilde{\boldsymbol{H}}^{(i)}(\boldsymbol{0})\tilde{\boldsymbol{y}} = \boldsymbol{A}^{(i)}\tilde{\boldsymbol{y}},\end{aligned}$$

which holds for an arbitrary $\tilde{\boldsymbol{y}} \in \mathring{C}^{(i)}$. Since $\mathring{C}^{(i)} \neq \emptyset$ by the definition of $\mathscr{I}'$, and $C^{(i)}$ is a closed convex, one has $\overline{\mathring{C}^{(i)}} = C^{(i)}$. Combining this with the continuity of the function $\tilde{\boldsymbol{y}} \mapsto \tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}})$, one arrives at (14).

Finally, let $\tilde{\boldsymbol{y}} \in \mathring{C}^{(i)} \cap C^{(j)}$, $i, j \in \mathscr{I}'$, $i \neq j$. In virtue of the equality $\overline{\mathring{C}^{(j)}} = C^{(j)}$, one can find $\tilde{\boldsymbol{z}} \in \mathring{C}^{(i)} \cap \mathring{C}^{(j)}$, and arguing as in (16), one gets $r > 0$ such that $r\tilde{\boldsymbol{z}} \in \mathring{\tilde{D}}^{(i)} \cap \mathring{\tilde{D}}^{(j)}$. This contradicts to (5). □

The previous theorem is completed by the following example, which shows that the indices from $\mathscr{I}(\bar{\boldsymbol{y}}) \setminus \mathscr{I}'$ not only may but even *have to* be omitted from calculation of $\tilde{\boldsymbol{H}}'$.

Example 1 Let $\boldsymbol{H}: \mathbb{R}^2 \to \mathbb{R}$ be given by

$$\begin{aligned}
&\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{H}^{(1)}(\boldsymbol{y}) = -y_1 - y_1^3 - y_2\\
&\quad \text{in } D^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ -y_1^3 - y_2 \le 0, y_1 \le 0\},\\
&\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{H}^{(2)}(\boldsymbol{y}) = -y_2\\
&\quad \text{in } D^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ -y_1 \le 0, -y_2 \le 0\},\\
&\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{H}^{(3)}(\boldsymbol{y}) = y_2\\
&\quad \text{in } D^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ y_2 \le 0, -y_1 \le 0\},\\
&\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{H}^{(4)}(\boldsymbol{y}) = -y_1 - y_1^3 + y_2\\
&\quad \text{in } D^{(4)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ y_1 \le 0, -y_1^3 + y_2 \le 0\},\\
&\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{H}^{(5)}(\boldsymbol{y}) = -y_1\\
&\quad \text{in } D^{(5)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ y_1^3 - y_2 \le 0, y_1^3 + y_2 \le 0\}
\end{aligned}$$

and take $\bar{\boldsymbol{y}} = \boldsymbol{0}$. Then

$$C^{(5)} = \{\boldsymbol{y} \in \mathbb{R}^2;\ y_2 = 0\}$$

but

$$\boldsymbol{H}'(\bar{\boldsymbol{y}};(1,0)) = 0 \neq -1 = \nabla\boldsymbol{H}^{(5)}(\bar{\boldsymbol{y}})(1,0). \qquad \square$$

Finally, we shall show that (6) captures correctly the essential features of the solution set of $\tilde{\boldsymbol{H}} = \boldsymbol{0}$ around $\boldsymbol{0}$, and so the one of $\boldsymbol{H} = \boldsymbol{0}$ around $\bar{\boldsymbol{y}}$. More precisely, we shall see that it determines completely the tangential directions of the solution curves of $(\mathscr{P})$ emanating from $\bar{\boldsymbol{y}}$. To this end, we shall need a non-degeneracy condition that prevents the solutions of (6) from being contained in the boundaries of the cones $C^{(i)}$, and consequently, the tangential directions of solution branches of $(\mathscr{P})$ from being tangent to the boundaries of the regions $D^{(i)}$. Such cases seem to occur rarely.

Assumption 3 *Let*

$$\operatorname{Ker}\nabla\boldsymbol{H}^{(i)}(\bar{\boldsymbol{y}}) \cap \Big(\bigcup_{j=1}^{M_i} \big\{\nabla G_j^{(i)}(\bar{\boldsymbol{y}})\big\}^{\perp}\Big) = \{\boldsymbol{0}\}, \quad \forall i \in \mathscr{I}(\bar{\boldsymbol{y}}).$$

Regarding [6], let us call the boundary solutions satisfying this condition *border-collision* solutions.

Theorem 2 *Let Assumptions 1, 2 and 3 hold and set*

$$\mathscr{I}'' := \{i \in \mathscr{I}(\bar{\boldsymbol{y}});\ \operatorname{Ker}\boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)} \neq \emptyset\}.$$

Then the solution set of $\tilde{\boldsymbol{H}}'(\boldsymbol{0};\tilde{\boldsymbol{y}}) = \boldsymbol{0}$ *coincides either with* $\{\boldsymbol{0}\}$ *if* $\mathscr{I}'' = \emptyset$*, or with the union* $\cup_{i\in\mathscr{I}''} \cup_{r\ge 0} r\tilde{\boldsymbol{y}}^{(i)}$*, where each* $\tilde{\boldsymbol{y}}^{(i)}$ *is arbitrarily chosen from* $\operatorname{Ker}\boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)}$.

In the first case, there exists a neighbourhood of $\mathbf{0}$ *such that the solution set of* $\tilde{\boldsymbol{H}} = \mathbf{0}$ *contains only* $\mathbf{0}$ *in it. In the second case, there are* $\delta^{(i)} > 0$ *and* C^1*-curves* $\boldsymbol{c}^{(i)}\colon [0, \delta^{(i)}) \to \mathbb{R}^{N+1}$, $i \in \mathscr{I}''$, *such that*

$$\left.\begin{aligned} &\text{(i) } \boldsymbol{c}^{(i)}(0) = \mathbf{0}, \quad \text{(ii) } \left(\boldsymbol{c}^{(i)}\right)'_+(0) \in \bigcup_{r>0} r\tilde{\boldsymbol{y}}^{(i)}, \\ &\text{(iii) } \forall s \in (0, \delta^{(i)})\colon \boldsymbol{c}^{(i)}(s) \in \mathring{\tilde{D}}^{(i)}, \end{aligned}\right\} \tag{17}$$

and the solution set of $\tilde{\boldsymbol{H}} = \mathbf{0}$ *coincides with the union* $\cup_{i\in\mathscr{I}''} \operatorname{Im}\boldsymbol{c}^{(i)}$ *in a vicinity of* $\mathbf{0}$.

Proof From Theorem 1 and Assumption 3, it is readily seen that $\tilde{\boldsymbol{y}}$ satisfies $\tilde{\boldsymbol{H}}'(\mathbf{0};\tilde{\boldsymbol{y}}) = \mathbf{0}$ iff $\tilde{\boldsymbol{y}} = \mathbf{0}$ or there exists $i \in \mathscr{I}'$ such that $\tilde{\boldsymbol{y}} \in \operatorname{Ker}\boldsymbol{A}^{(i)} \cap \mathring{\tilde{C}}^{(i)}$. The first part of the assertion then follows immediately from Assumption 2 and (12).

For the second part, Assumption 1 guarantees that the solution set $\{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) = \mathbf{0}\}$ is contained in $\cup_{i\in\mathscr{I}(\tilde{\boldsymbol{y}})}\{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}^{(i)}(\tilde{\boldsymbol{y}}) = \mathbf{0}\}$. So, consider $i \in \mathscr{I}(\tilde{\boldsymbol{y}})$ fixed.

In virtue of Assumption 2, N columns of $\boldsymbol{\nabla}\tilde{\boldsymbol{H}}^{(i)}(\mathbf{0})$ are linearly independent. We shall consider for definiteness the case when these are the first N columns, that is, $\boldsymbol{\nabla}_{(\tilde{y}_1,\dots,\tilde{y}_N)}\tilde{\boldsymbol{H}}^{(i)}(\mathbf{0})$ is non-singular, the other cases being analogous. According to the implicit-function theorem, $\tilde{\boldsymbol{H}}^{(i)} = \mathbf{0}$ determines a unique implicit function $\tilde{y}_{N+1} \mapsto (\tilde{y}_1(\tilde{y}_{N+1}), \dots, \tilde{y}_N(\tilde{y}_{N+1}))$ in a vicinity of $\mathbf{0}$. Defining,

$$\boldsymbol{c}\colon s \mapsto (\tilde{y}_1(s), \dots, \tilde{y}_N(s), s),$$

one gets from the chain rule that

$$\mathbf{0} = \boldsymbol{\nabla}\tilde{\boldsymbol{H}}^{(i)}(\boldsymbol{c}(0))\boldsymbol{c}'(0) = \boldsymbol{\nabla}\tilde{\boldsymbol{H}}^{(i)}(\mathbf{0})\boldsymbol{c}'(0),$$

that is, $\boldsymbol{c}'(0) \in \operatorname{Ker}\boldsymbol{\nabla}\tilde{\boldsymbol{H}}^{(i)}(\mathbf{0}) = \operatorname{Ker}\boldsymbol{A}^{(i)}$. By Assumption 3,

$$\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) \neq 0, \quad j = 1, \dots, M_i.$$

Take j fixed and consider the case $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) < 0$. There exist $\varepsilon > 0$ and $\eta > 0$ such that

$$\forall \tilde{\boldsymbol{y}} \in B(\boldsymbol{c}'(0), \eta)\colon \quad \nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\tilde{\boldsymbol{y}} < -\varepsilon, \tag{18}$$

where $B(\boldsymbol{c}'(0), \eta)$ stands for the closed ball centred at $\boldsymbol{c}'(0)$ with the radius η. The mean-value theorem gives for any $r > 0$ sufficiently small,

$$\begin{aligned} &|\tilde{G}_j^{(i)}(r\tilde{\boldsymbol{y}}) - \tilde{G}_j^{(i)}(\mathbf{0}) - \nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})r\tilde{\boldsymbol{y}}| \\ &\qquad \leq \sup_{t\in[0,1]} \|\nabla \tilde{G}_j^{(i)}(tr\tilde{\boldsymbol{y}}) - \nabla \tilde{G}_j^{(i)}(\mathbf{0})\|\|r\tilde{\boldsymbol{y}}\|. \end{aligned} \tag{19}$$

Denoting

$$M := \max\{\|\tilde{\boldsymbol{y}}\|;\ \tilde{\boldsymbol{y}} \in B(\boldsymbol{c}'(0), \eta)\},$$

one can find $\eta_1 > 0$ such that

$$\forall \tilde{\boldsymbol{z}} \in B(\mathbf{0}, \eta_1)\colon \quad \|\nabla \tilde{G}_j^{(i)}(\tilde{\boldsymbol{z}}) - \nabla \tilde{G}_j^{(i)}(\mathbf{0})\| \leq \frac{\varepsilon}{M}. \tag{20}$$

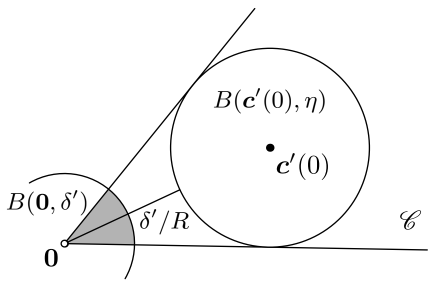


Fig. 2 The intersection $\mathscr{C} \cap B(\mathbf{0}, \delta')$.

Combining (19) with (2) and (20), one obtains for $R := \eta_1/M$ that

$$\begin{aligned} &\forall r \in (0, R)\ \forall \tilde{\boldsymbol{y}} \in B(\boldsymbol{c}'(0), \eta)\colon \\ &\qquad r\tilde{\boldsymbol{y}} \in B(\mathbf{0}, \eta_1), \\ &\qquad |\tilde{G}_j^{(i)}(r\tilde{\boldsymbol{y}}) - \nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})r\tilde{\boldsymbol{y}}| \leq r\varepsilon. \end{aligned}$$

This together with (18) yields

$$\forall r \in (0, R)\ \forall \tilde{\boldsymbol{y}} \in B(\boldsymbol{c}'(0), \eta)\colon \quad \tilde{G}_j^{(i)}(r\tilde{\boldsymbol{y}}) < 0. \tag{21}$$

Next, introduce a cone $\mathscr{C}$ and a real δ' by

$$\begin{aligned} &\mathscr{C} := \bigcup_{r>0} rB(\boldsymbol{c}'(0), \eta), \\ &\delta' := R\min\{\|\tilde{\boldsymbol{y}}\|;\ \tilde{\boldsymbol{y}} \in B(\boldsymbol{c}'(0), \eta)\} \quad (> 0 \text{ by (18)}). \end{aligned}$$

It is readily seen from (21) and Fig. 2 that

$$\forall \tilde{\boldsymbol{y}} \in \mathscr{C} \cap B(\mathbf{0}, \delta')\colon \quad \tilde{G}_j^{(i)}(\tilde{\boldsymbol{y}}) < 0. \tag{22}$$

Making use of the differentiability and continuity of $\boldsymbol{c}$, one can take $\delta_j > 0$ sufficiently small so that

$$\begin{aligned} \forall s \in (0, \delta_j)\colon \quad &\frac{\boldsymbol{c}(s) - \boldsymbol{c}(0)}{s} \in B(\boldsymbol{c}'(0), \eta), \\ &\boldsymbol{c}(s) - \boldsymbol{c}(0) \in B(\mathbf{0}, \delta'). \end{aligned}$$

Inserting $\boldsymbol{c}(0) = \mathbf{0}$, one can see from here that

$$\forall s \in (0, \delta_j)\colon \quad \boldsymbol{c}(s) \in \mathscr{C} \cap B(\mathbf{0}, \delta'),$$

and (22) leads to

$$\forall s \in (0, \delta_j)\colon \quad \tilde{G}_j^{(i)}(\boldsymbol{c}(s)) < 0.$$

Reducing δ_j if necessary, one can show by analogous reasoning that

$$\forall s \in (-\delta_j, 0)\colon \quad \tilde{G}_j^{(i)}(\boldsymbol{c}(s)) > 0.$$

Clearly, the last two inequalities are reversed in the case $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) > 0$. Repeating the argumentation for all indices j, one gets $\delta^{(i)} > 0$ guaranteeing that for any $j \in \{1, \dots,$

$M_i\}$,

$\tilde{G}_j^{(i)}(\boldsymbol{c}(s)) > 0,\ \forall s \in (-\delta^{(i)}, 0)$,

and $\tilde{G}_j^{(i)}(\boldsymbol{c}(s)) < 0,\ \forall s \in (0, \delta^{(i)})$, if $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) < 0$,

$\tilde{G}_j^{(i)}(\boldsymbol{c}(s)) < 0,\ \forall s \in (-\delta^{(i)}, 0)$,

and $\tilde{G}_j^{(i)}(\boldsymbol{c}(s)) > 0,\ \forall s \in (0, \delta^{(i)})$, if $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) > 0$.

Three different cases may occur for the index i still kept fixed:

1. $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) < 0$ for any j. In this case,

$$\boldsymbol{c}'(0) \in \operatorname{Ker} \boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)},$$
$$\boldsymbol{c}(s) \in \mathring{\tilde{D}}^{(i)} \text{ for any } s \in (0, \delta^{(i)}),$$
$$\boldsymbol{c}(s) \notin \tilde{D}^{(i)} \text{ for any } s \in (-\delta^{(i)}, 0).$$

Hence, $i \in \mathscr{I}''$,

$$\boldsymbol{c}'(0) \in \bigcup_{r>0} r\tilde{\boldsymbol{y}}^{(i)}$$

(with $\tilde{\boldsymbol{y}}^{(i)}$ from the first part of the assertion), and (17) is satisfied for $\boldsymbol{c}^{(i)} := \boldsymbol{c}_{|[0,\delta^{(i)})}$. One obtains immediately that $\{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) = \mathbf{0}\} \cap \{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}^{(i)}(\tilde{\boldsymbol{y}}) = \mathbf{0}\}$ coincides with $\operatorname{Im} \boldsymbol{c}^{(i)}$ in a vicinity of $\mathbf{0}$.

2. $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) > 0$ for any j. In an analogous way, (17) is satisfied for $\boldsymbol{c}^{(i)} \colon [0, \delta^{(i)}) \ni s \mapsto \boldsymbol{c}(-s)$ and $\{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) = \mathbf{0}\} \cap \{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}^{(i)}(\tilde{\boldsymbol{y}}) = \mathbf{0}\}$ coincides with $\operatorname{Im} \boldsymbol{c}^{(i)}$ in a vicinity of $\mathbf{0}$.

3. $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) < 0$ for some j and $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\boldsymbol{c}'(0) > 0$ for the others j. In this case,

$$\boldsymbol{c}'(0) \notin C^{(i)}, \quad \boldsymbol{c}(s) \notin \tilde{D}^{(i)} \text{ for any } s \in (-\delta^{(i)}, 0) \cup (0, \delta^{(i)}),$$

$i \notin \mathscr{I}''$ and $\{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}(\tilde{\boldsymbol{y}}) = \mathbf{0}\} \cap \{\tilde{\boldsymbol{y}} \in \tilde{O};\ \tilde{\boldsymbol{H}}^{(i)}(\tilde{\boldsymbol{y}}) = \mathbf{0}\}$ shrinks to $\{\mathbf{0}\}$ in a vicinity of $\mathbf{0}$.

The second part of the theorem is then proved by getting together the respective cases for all $i \in \mathscr{I}(\tilde{\boldsymbol{y}})$. □

To summarise, we arrive at the following simplified system for branching in a border-collision solution of $(\mathscr{P})$, the so-called border-collision *normal form* (omitting the tildes for brevity of notation in what follows):

$$\left.\begin{array}{l} \textit{Find } \boldsymbol{y} \in \mathbb{R}^{N+1} \textit{ such that} \\ \boldsymbol{F}(\boldsymbol{y}) = \mathbf{0}, \end{array}\right\} \tag{NF}$$

where $\boldsymbol{F} \colon \mathbb{R}^{N+1} \to \mathbb{R}^N$ is a piecewise-linear function defined by

$$\boldsymbol{F}(\boldsymbol{y}) := \boldsymbol{A}^{(i)}\boldsymbol{y}, \quad \boldsymbol{y} \in C^{(i)},\ i \in \mathscr{I}', \tag{23}$$

with $\boldsymbol{A}^{(i)}$, $C^{(i)}$ and $\mathscr{I}'$ introduced in (7)–(10).

It is worth noticing that (6) is exactly the first-order system $(\mathscr{P}')$ from [21, Subsection 2.1] shifted to $\mathbf{0}$. The relation between $(\mathscr{P})$ and $(\mathscr{P}')$ was studied for a general PC^1-function $\boldsymbol{H}$. Among others, scenarios resulting from violation of Assumption 3 were shown in op.cit., Example 1. Some criteria guaranteeing the existence and uniqueness of solutions of $(\mathscr{P}')$ can be found in that paper and in references therein either. Nevertheless, it seems to be still difficult to prove a general assertion determining completely the structure of solutions of $(\mathscr{P}')$ or of its special case (NF) although these are already simplifications of $(\mathscr{P})$. That is why we shall investigate more closely the simplest cases of (NF), which are most likely to occur.

2.2 The Simplest Branching Scenarios

In this subsection, we shall consider Problem (NF) with a general piecewise-linear function $\boldsymbol{F}$ defined by (23), where $\mathscr{I}' = \{1, \dots, L\}$ for some $L \in \mathbb{N}$, $\boldsymbol{A}^{(i)} \in \mathbb{R}^{N \times (N+1)}$ are arbitrary matrices and $C^{(i)}$ are given by (9) with arbitrary matrices $\boldsymbol{B}^{(i)} \in \mathbb{R}^{M_i \times (N+1)}$. In accordance with Assumptions 2 and 3 and Theorem 1, we shall suppose that (11)–(13) hold and

$$\mathring{C}^{(i)} \neq \emptyset, \quad \operatorname{rank} \boldsymbol{A}^{(i)} = N, \quad \forall i \in \mathscr{I}',$$
$$\operatorname{Ker} \boldsymbol{A}^{(i)} \cap \Big(\bigcup_{j=1}^{M_i} \{\boldsymbol{B}_j^{(i)\top}\}^\perp \Big) = \{\mathbf{0}\}, \quad \forall i \in \mathscr{I}', \tag{24}$$

where $\boldsymbol{B}_j^{(i)}$ stands for the jth row vector of $\boldsymbol{B}^{(i)}$.

We shall examine possible branching scenarios in the cases with $L \in \{2,3,4\}$ and $M_i \in \{1,2\}$. Moreover, introducing the block matrix $\boldsymbol{B} \in \mathbb{R}^{(M_1 + \dots + M_L) \times (N+1)}$ as

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}^{(1)} \\ \vdots \\ \boldsymbol{B}^{(L)} \end{pmatrix},$$

we shall restrict ourselves to the cases with $\operatorname{rank} \boldsymbol{B} \in \{1,2\}$. In these cases, $C^{(i)}$, $i \in \mathscr{I}'$, can be represented by their projections into a two-dimensional space. Indeed, let V be any two-dimensional space containing $\operatorname{Im} \boldsymbol{B}^\top$ if $\operatorname{rank} \boldsymbol{B} = 1$, and $V = \operatorname{Im} \boldsymbol{B}^\top$ if $\operatorname{rank} \boldsymbol{B} = 2$. Then $\boldsymbol{y}$ lies in $C^{(i)}$ if and only if $P\boldsymbol{y}$ lies in $PC^{(i)}$, where P denotes the orthogonal projection onto V, and

$$PC^{(i)} = \{\boldsymbol{z} \in V;\ \boldsymbol{B}^{(i)}\boldsymbol{z} \le \mathbf{0}\}.$$

This will simplify our considerations. In particular, omitting superfluous inequalities if necessary, we may assume that $\operatorname{rank} \boldsymbol{B}^{(i)} = M_i$.

Our study will be based on the following elementary result:

Lemma 1 *Let $\boldsymbol{F}: \mathbb{R}^{N+1} \to \mathbb{R}^n$ be a piecewise-linear function with two selections such that*

$$\boldsymbol{F}(\boldsymbol{y}) = \boldsymbol{A}^{(i)}\boldsymbol{y}, \quad \boldsymbol{y} \in C^{(i)},\ i = 1,2,$$
$$C^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}^\top \boldsymbol{y} \le 0\}, \quad C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}^\top \boldsymbol{y} \ge 0\}$$

for some $\boldsymbol{A}^{(i)} \in \mathbb{R}^{N\times(N+1)}$ and $\boldsymbol{b} \in \mathbb{R}^{N+1}$ with

$$\operatorname{rank}\boldsymbol{A}^{(i)} = N, \quad \operatorname{Ker}\boldsymbol{A}^{(i)} \cap \{\boldsymbol{b}\}^\perp = \{\boldsymbol{0}\}, \quad i = 1,2. \tag{25}$$

Then any two vectors $\boldsymbol{y}^{(i)} \in \operatorname{Ker}\boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)}$, $i = 1,2$, satisfy

$$\det\begin{pmatrix}\boldsymbol{A}^{(1)}\\ -\boldsymbol{y}^{(1)\top}\end{pmatrix}\det\begin{pmatrix}\boldsymbol{A}^{(2)}\\ \boldsymbol{y}^{(2)\top}\end{pmatrix} > 0.$$

Proof Since $\boldsymbol{b} \neq \boldsymbol{0}$ by (25), there exists an orthonormal basis $\{\boldsymbol{q}_1, \dots \boldsymbol{q}_{N+1}\}$ of $\mathbb{R}^{N+1}$ with $\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, and one can introduce a coordinate transformation $\boldsymbol{T}: \mathbb{R}^{N+1} \to \mathbb{R}^{N+1}$ by

$$\boldsymbol{T}: \boldsymbol{y} \mapsto \hat{\boldsymbol{y}} := \boldsymbol{Q}\boldsymbol{y}$$

with a matrix $\boldsymbol{Q}$ whose rows are formed by $\boldsymbol{q}_1^\top, \dots, \boldsymbol{q}_{N+1}^\top$. Then, one can define a piecewise-linear function $\hat{\boldsymbol{F}}: \mathbb{R}^{N+1} \to \mathbb{R}^N$ by

$$\hat{\boldsymbol{F}}(\hat{\boldsymbol{y}}) := \boldsymbol{F}(\boldsymbol{T}^{-1}(\hat{\boldsymbol{y}})) = \boldsymbol{F}(\boldsymbol{Q}^\top \hat{\boldsymbol{y}})$$

with selections $\hat{\boldsymbol{F}}^{(i)} = \boldsymbol{F}^{(i)} \circ \boldsymbol{T}^{-1}$ and regions $\hat{C}^{(i)} = \boldsymbol{T}(C^{(i)})$, $i = 1,2$. Clearly,

$$\hat{C}^{(1)} = \{\hat{\boldsymbol{y}} \in \mathbb{R}^{N+1};\ \hat{y}_1 \le 0\}, \quad \hat{C}^{(2)} = \{\hat{\boldsymbol{y}} \in \mathbb{R}^{N+1};\ \hat{y}_1 \ge 0\},$$

and one can write

$$\hat{\boldsymbol{F}}(\hat{\boldsymbol{y}}) = \begin{cases}\hat{\boldsymbol{A}}_1^{(1)}\hat{y}_1 + \hat{\boldsymbol{A}}_2^{(1)}\hat{\boldsymbol{y}}_2 & \text{if } \hat{y}_1 \le 0,\\ \hat{\boldsymbol{A}}_1^{(2)}\hat{y}_1 + \hat{\boldsymbol{A}}_2^{(2)}\hat{\boldsymbol{y}}_2 & \text{if } \hat{y}_1 \ge 0\end{cases}$$

with $\hat{\boldsymbol{A}}_1^{(i)} \in \mathbb{R}^N$, $\hat{\boldsymbol{A}}_2^{(i)} \in \mathbb{R}^{N\times N}$, $(\hat{\boldsymbol{A}}_1^{(i)},\ \hat{\boldsymbol{A}}_2^{(i)}) := \boldsymbol{A}^{(i)}\boldsymbol{Q}^\top$ and $\hat{\boldsymbol{y}}_2 := (\hat{y}_2, \dots, \hat{y}_{N+1})$.

As $\hat{\boldsymbol{F}}$ is continuous, we have $\hat{\boldsymbol{A}}_2^{(1)} = \hat{\boldsymbol{A}}_2^{(2)}$, and we shall denote it simply by $\hat{\boldsymbol{A}}_2$. Moreover, we claim that it is non-singular. In view of

$$\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(1)} & \hat{\boldsymbol{A}}_2\\ 1 & \boldsymbol{0}\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}^{(1)}\\ \boldsymbol{q}_1^\top\end{pmatrix}\boldsymbol{Q}^\top,$$

it suffices to show that the matrix $\begin{pmatrix}\boldsymbol{A}^{(1)}\\ \boldsymbol{q}_1^\top\end{pmatrix}$ is non-singular. But from the imposed assumptions, $\operatorname{Ker}\boldsymbol{A}^{(1)}$ is spanned by $\boldsymbol{y}^{(1)}$ and $\boldsymbol{y}^{(1)} \in \mathring{C}^{(1)}$, which entails that $\boldsymbol{q}_1^\top \boldsymbol{y}^{(1)} < 0$. Thus, $\boldsymbol{q}_1 \notin (\operatorname{Ker}\boldsymbol{A}^{(1)})^\perp = \operatorname{Im}\boldsymbol{A}^{(1)\top}$. Taking into account that $\operatorname{rank}\boldsymbol{A}^{(1)}$ equals N, one deduces that $\operatorname{rank}(\boldsymbol{A}^{(1)\top}\ \boldsymbol{q}_1)$ must equal $N+1$, that is, $(\boldsymbol{A}^{(1)\top}\ \boldsymbol{q}_1)$ is non-singular and so are its transpose as well as $\hat{\boldsymbol{A}}_2$.

Now, if one takes $(\hat{y}_1^{(i)}, \hat{\boldsymbol{y}}_2^{(i)}) := \boldsymbol{T}(\boldsymbol{y}^{(i)}) = \boldsymbol{Q}\boldsymbol{y}^{(i)}$, then $\hat{\boldsymbol{A}}_1^{(i)}\hat{y}_1^{(i)} + \hat{\boldsymbol{A}}_2\hat{\boldsymbol{y}}_2^{(i)} = \boldsymbol{0}$, which gives

$$\hat{\boldsymbol{y}}_2^{(i)} = -\hat{y}_1^{(i)}\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(i)}.$$

Here, $\hat{y}_1^{(1)} < 0$ and $\hat{y}_1^{(2)} > 0$ as $\hat{\boldsymbol{y}}^{(1)}$ and $\hat{\boldsymbol{y}}^{(2)}$ are from $\mathring{\hat{C}}^{(1)}$ and $\mathring{\hat{C}}^{(2)}$, respectively. This yields finally:

$$\begin{aligned}
&\det\begin{pmatrix}\boldsymbol{A}^{(1)}\\ -\boldsymbol{y}^{(1)\top}\end{pmatrix}\det\begin{pmatrix}\boldsymbol{A}^{(2)}\\ \boldsymbol{y}^{(2)\top}\end{pmatrix}\\
&\quad= \det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(1)} & \hat{\boldsymbol{A}}_2\\ -\hat{y}_1^{(1)} & -\hat{\boldsymbol{y}}_2^{(1)\top}\end{pmatrix}\det\boldsymbol{Q}\det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(2)} & \hat{\boldsymbol{A}}_2\\ \hat{y}_1^{(2)} & \hat{\boldsymbol{y}}_2^{(2)\top}\end{pmatrix}\det\boldsymbol{Q}\\
&\quad= \left(-\hat{y}_1^{(1)}\right)\det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(1)} & \hat{\boldsymbol{A}}_2\\ 1 & -\left(\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(1)}\right)^\top\end{pmatrix}\\
&\qquad\cdot\hat{y}_1^{(2)}\det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(2)} & \hat{\boldsymbol{A}}_2\\ 1 & -\left(\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(2)}\right)^\top\end{pmatrix}\\
&\quad> 0
\end{aligned}$$

because

$$\begin{aligned}
&\det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(i)} & \hat{\boldsymbol{A}}_2\\ 1 & -\left(\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(i)}\right)^\top\end{pmatrix}\\
&\quad= \det\begin{pmatrix}\hat{\boldsymbol{A}}_1^{(i)} & \hat{\boldsymbol{A}}_2\\ 1 + \left(\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(i)}\right)^\top\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(i)} & \boldsymbol{0}\end{pmatrix}\\
&\quad= (-1)^{N+2}\left(1 + \left\|\hat{\boldsymbol{A}}_2^{-1}\hat{\boldsymbol{A}}_1^{(i)}\right\|^2\right)\det\hat{\boldsymbol{A}}_2.
\end{aligned}$$

□

Let us now describe all possible cases of (NF) under the restrictions introduced above.

I. $L = 2$.

Invoking that $C^{(1)}$ and $C^{(2)}$ are closed convex polyhedral cones satisfying (11)–(13), one can see that necessarily $M_1 = M_2 = 1$, and

$$C^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}^\top \boldsymbol{y} \le 0\}, \quad C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}^\top \boldsymbol{y} \ge 0\}$$

for some $\boldsymbol{0} \neq \boldsymbol{b} \in \mathbb{R}^{N+1}$. As $\boldsymbol{b}^\top \boldsymbol{y} \neq 0$ for any $\boldsymbol{y} \in \operatorname{Ker}\boldsymbol{A}^{(i)} \setminus \{\boldsymbol{0}\}$ by (24), one can find $\boldsymbol{y}^{(i)} \in \operatorname{Ker}\boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)}$, $i = 1,2$, and the solution set of (NF) consists of two rays generated by $\boldsymbol{y}^{(1)}$ and $\boldsymbol{y}^{(2)}$:

$$\bigcup_{i=1}^{2}\bigcup_{r\ge 0} r\boldsymbol{y}^{(i)}$$

(Fig. 3). Furthermore, direct application of Lemma 1 gives:

$$\det\begin{pmatrix}\boldsymbol{A}^{(1)}\\ -\boldsymbol{y}^{(1)\top}\end{pmatrix}\det\begin{pmatrix}\boldsymbol{A}^{(2)}\\ \boldsymbol{y}^{(2)\top}\end{pmatrix} > 0.$$

If this condition holds, we shall say that the solution rays are *coherently oriented*, if the converse inequality holds, we shall speak about *incoherent orientation*.

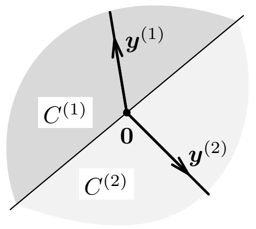


Fig. 3 Solution set in Case I.

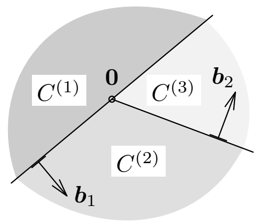


Fig. 4 Structure of the regions in Case II.

A simple example (with $N = 1$):

$$F(\boldsymbol{y}) = (y_1)_+ + y_2 = \begin{cases} y_2 & \text{if } y_1 \le 0, \\ y_1 + y_2 & \text{if } y_1 \ge 0. \end{cases}$$

II. $L = 3, M_1 = 1$.
One can see that the only possibility under the imposed restrictions is: $M_2 = M_3 = 2$, and

$$C^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0\},$$
$$C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0, \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\},$$
$$C^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0, \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0\}$$

for some linearly independent vectors $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathbb{R}^{N+1}$ (see Fig. 4). Due to the continuity of $\boldsymbol{F}$,

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(2)} \text{ in } C^{(1)} \cap C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} = 0, \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\},$$

and the linearity of $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$ entails

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(2)} \text{ in } \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} = 0\} = \{\boldsymbol{b}_1\}^\perp.$$

By analogous argumentation, one gets

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(3)} \text{ in } \{\boldsymbol{b}_1\}^\perp,$$

and consequently, $\boldsymbol{F}^{(2)} = \boldsymbol{F}^{(3)}$ in $\{\boldsymbol{b}_1\}^\perp$. On the other hand, one can deduce also that $\boldsymbol{F}^{(2)} = \boldsymbol{F}^{(3)}$ in $\{\boldsymbol{b}_2\}^\perp$. From here, one obtains that

$$\boldsymbol{F}^{(2)} = \boldsymbol{F}^{(3)} \text{ in } \operatorname{span}(\{\boldsymbol{b}_1\}^\perp \cup \{\boldsymbol{b}_2\}^\perp) = \mathbb{R}^{N+1}$$

because $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are linearly independent. This means that there are only two distinct selections, in fact, and one recovers exactly the situation from Case I.

III. $L = 3, M_i = 2, i = 1, 2, 3$.
In this case, one can write

$$C^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0, \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\},$$
$$C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0, \boldsymbol{b}_3^\top \boldsymbol{y} \le 0\},$$
$$C^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0, \boldsymbol{b}_3^\top \boldsymbol{y} \ge 0\},$$

where any two of the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3 \in \mathbb{R}^{N+1}$ are linearly independent, and

$$C^{(1)} \cap C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0, \boldsymbol{b}_2^\top \boldsymbol{y} = 0\}, \tag{26}$$
$$C^{(2)} \cap C^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0, \boldsymbol{b}_3^\top \boldsymbol{y} = 0\},$$
$$C^{(1)} \cap C^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} = 0, \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\}$$

by the convexity of each $C^{(i)}$ (Fig. 5(a)).

First, we shall show that (NF) cannot have three distinct solution rays in this case: Suppose for contradiction that

$$\boldsymbol{y}^{(i)} \in \operatorname{Ker} \boldsymbol{A}^{(i)} \cap \mathring{C}^{(i)}, \quad i = 1, 2, 3,$$

taking into account (24). Owing to the continuity of $\boldsymbol{F}$ and (26),

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(2)} \text{ in } C^{(1)} \cap C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0, \boldsymbol{b}_2^\top \boldsymbol{y} = 0\},$$

and the linearity of $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$ implies that

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(2)} \text{ in } \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} = 0\}.$$

Hence, defining

$$C^{(12)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\}, \quad C^{(21)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0\},$$

one obtains

$$\boldsymbol{F}^{(1)} = \boldsymbol{F}^{(2)} \text{ in } C^{(12)} \cap C^{(21)}, \quad \boldsymbol{y}^{(1)} \in \mathring{C}^{(12)}, \quad \boldsymbol{y}^{(2)} \in \mathring{C}^{(21)}.$$

It follows from Lemma 1 when applied to the piecewise-linear function defined to be $\boldsymbol{F}^{(1)}$ in $C^{(12)}$ and $\boldsymbol{F}^{(2)}$ in $C^{(21)}$ that

$$\det \begin{pmatrix} \boldsymbol{A}^{(1)} \\ -\boldsymbol{y}^{(1)\top} \end{pmatrix} \det \begin{pmatrix} \boldsymbol{A}^{(2)} \\ \boldsymbol{y}^{(2)\top} \end{pmatrix} > 0. \tag{27}$$

By similar reasoning, one also gets

$$\det \begin{pmatrix} \boldsymbol{A}^{(3)} \\ -\boldsymbol{y}^{(3)\top} \end{pmatrix} \det \begin{pmatrix} \boldsymbol{A}^{(2)} \\ \boldsymbol{y}^{(2)\top} \end{pmatrix} > 0, \tag{28}$$
$$\det \begin{pmatrix} \boldsymbol{A}^{(3)} \\ -\boldsymbol{y}^{(3)\top} \end{pmatrix} \det \begin{pmatrix} \boldsymbol{A}^{(1)} \\ \boldsymbol{y}^{(1)\top} \end{pmatrix} > 0, \tag{29}$$

and the product of (27), (28) and (29) leads to a contradiction.

One can exclude the scenario with exactly one solution ray, as well: For definiteness, suppose that

$$\boldsymbol{y}^{(1)} \in \operatorname{Ker} \boldsymbol{A}^{(1)} \cap \mathring{C}^{(1)}, \quad \operatorname{Ker} \boldsymbol{A}^{(i)} \cap C^{(i)} = \{\mathbf{0}\},\ i = 2, 3.$$

Making use of (24), one can find $\boldsymbol{y}^{(2)}$ and $\boldsymbol{y}^{(3)}$ such that

$$\boldsymbol{y}^{(2)} \in \operatorname{Ker} \boldsymbol{A}^{(2)}, \quad \boldsymbol{b}_2^\top \boldsymbol{y}^{(2)} > 0, \quad \boldsymbol{b}_3^\top \boldsymbol{y}^{(3)} > 0,$$
$$\boldsymbol{y}^{(3)} \in \operatorname{Ker} \boldsymbol{A}^{(3)}, \quad \boldsymbol{b}_3^\top \boldsymbol{y}^{(3)} > 0, \quad \boldsymbol{b}_1^\top \boldsymbol{y}^{(3)} < 0.$$

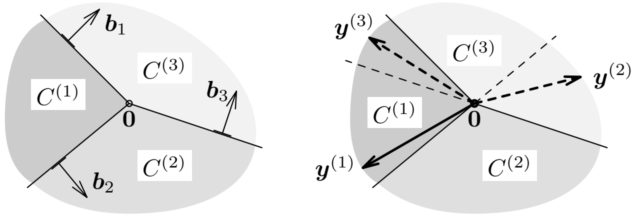


Fig. 5 Case III: (a) structure of the regions; (b) scenario with one solution ray.

(the so-called *virtual solutions*; see Fig. 5(b)). Taking $C^{(12)}$ and $C^{(21)}$ as before, one recovers (27) by the same reasoning. Further, defining

$$C^{(23)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_3^\top \boldsymbol{y} \le 0\}, \quad C^{(32)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_3^\top \boldsymbol{y} \ge 0\},$$

and using Lemma 1 for the piecewise-linear function introduced as $\boldsymbol{F}^{(2)}$ in $C^{(23)}$ and $\boldsymbol{F}^{(3)}$ in $C^{(32)}$ with $-\boldsymbol{y}^{(2)} \in \mathring{C}^{(23)}$ and $\boldsymbol{y}^{(3)} \in \mathring{C}^{(32)}$, one deduces that

$$\det\begin{pmatrix}\boldsymbol{A}^{(2)}\\ \boldsymbol{y}^{(2)\top}\end{pmatrix}\det\begin{pmatrix}\boldsymbol{A}^{(3)}\\ \boldsymbol{y}^{(3)\top}\end{pmatrix} > 0. \tag{30}$$

Eventually, one gets in a similar way that

$$\det\begin{pmatrix}\boldsymbol{A}^{(1)}\\ -\boldsymbol{y}^{(1)\top}\end{pmatrix}\det\begin{pmatrix}\boldsymbol{A}^{(3)}\\ -\boldsymbol{y}^{(3)\top}\end{pmatrix} > 0. \tag{31}$$

Comparing (27), (30) and (31), one arrives at a contradiction.

On the other hand, examining the scenario with two distinct solution rays, one can conclude that it is admissible, the two rays being *coherently oriented*. The second admissible scenario is that (NF) has only a trivial solution. Simple examples of these two scenarios follow.

(i) Trivial solution set:

$$F(\boldsymbol{y}) = y_1 + y_2 - 3P_{[(y_1)_+,+\infty)}(y_2) = \begin{cases} y_1 + y_2 & \text{if } y_1 \le 0,\, y_2 \le 0,\\ y_2 - 2y_1 & \text{if } y_1 \ge 0,\, y_1 \ge y_2,\\ y_1 - 2y_2 & \text{if } y_2 \ge 0,\, y_2 \ge y_1.\end{cases}$$

(ii) Two solution rays:

$$F(\boldsymbol{y}) = y_1 + y_2 + P_{[(y_1)_+,+\infty)}(y_2) = \begin{cases} y_1 + y_2 & \text{if } y_1 \le 0,\, y_2 \le 0,\\ 2y_1 + y_2 & \text{if } y_1 \ge 0,\, y_1 \ge y_2,\\ y_1 + 2y_2 & \text{if } y_2 \ge 0,\, y_2 \ge y_1.\end{cases}$$

IV. $L = 4$, $M_1 = 1$.
The only possible situation under our considerations is: $M_i = 2$, $i = 2,3,4$, and

$$C^{(1)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0\},$$
$$C^{(2)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0,\, \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\},$$
$$C^{(3)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0,\, \boldsymbol{b}_3^\top \boldsymbol{y} \le 0\},$$
$$C^{(4)} = \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0,\, \boldsymbol{b}_3^\top \boldsymbol{y} \ge 0\}$$

for some $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3 \in \mathbb{R}^{N+1}$ such that any two of them are linearly independent and

$$C^{(3)} \subset \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0\}$$

(Fig. 6).

Thanks to (24), the intersection $\operatorname{Ker}\boldsymbol{A}^{(1)} \cap \mathring{C}^{(1)}$ is always non-empty, that is, (NF) has at least one solution ray. One can show in a similar way as in Case III that the scenarios with exactly one and three solution rays are not possible. The solution set consists of either two or four solution rays. If there are only two rays, they are always *coherently oriented*. If there are four rays, the rays in any two *adjacent* cones are *coherently* oriented whereas the rays in any two *opposite* cones are *incoherently* oriented.

Examples:

(i) Solution rays in two adjacent cones:

$$F(\boldsymbol{y}) = 4y_1 + y_2 + 2P_{[-(y_1)_+,(y_1)_+]}(y_2) = \begin{cases} 4y_1 + y_2 & \text{if } y_1 \le 0,\\ 4y_1 + 3y_2 & \text{if } |y_2| \le y_1,\\ 2y_1 + y_2 & \text{if } y_2 \le -y_1 \le 0,\\ 6y_1 + y_2 & \text{if } y_2 \ge y_1 \ge 0.\end{cases}$$

(ii) Solution rays in two opposite cones:

$$F(\boldsymbol{y}) = y_2 + 2P_{[-(y_1)_+,(y_1)_+]}(y_2) = \begin{cases} y_2 & \text{if } y_1 \le 0,\\ 3y_2 & \text{if } |y_2| \le y_1,\\ y_2 - 2y_1 & \text{if } y_2 \le -y_1 \le 0,\\ y_2 + 2y_1 & \text{if } y_2 \ge y_1 \ge 0.\end{cases}$$

(iii) Four solution rays:

$$F(\boldsymbol{y}) = -y_2 + 2P_{[-(y_1)_+,(y_1)_+]}(y_2) = \begin{cases} -y_2 & \text{if } y_1 \le 0,\\ y_2 & \text{if } |y_2| \le y_1,\\ -y_2 - 2y_1 & \text{if } y_2 \le -y_1 \le 0,\\ -y_2 + 2y_1 & \text{if } y_2 \ge y_1 \ge 0.\end{cases}$$

V. $L = 4$, $M_i = 2$, $i = 1,2,3,4$.
One can write

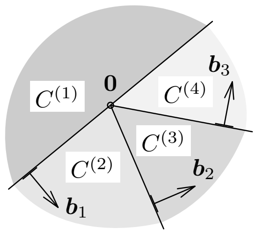


Fig. 6 Case IV.

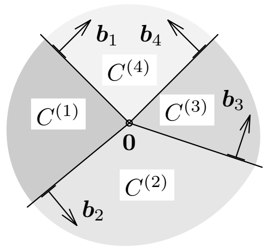


Fig. 7 Case V.

$$
\begin{aligned}
C^{(1)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0,\ \boldsymbol{b}_2^\top \boldsymbol{y} \le 0\},\\
C^{(2)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0,\ \boldsymbol{b}_3^\top \boldsymbol{y} \le 0\},\\
C^{(3)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_3^\top \boldsymbol{y} \ge 0,\ \boldsymbol{b}_4^\top \boldsymbol{y} \le 0\},\\
C^{(4)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \ge 0,\ \boldsymbol{b}_4^\top \boldsymbol{y} \ge 0\}
\end{aligned}
$$

with $\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3,\boldsymbol{b}_4 \in \mathbb{R}^{N+1}$ such that the couples $\{\boldsymbol{b}_1,\boldsymbol{b}_2\}$, $\{\boldsymbol{b}_2,\boldsymbol{b}_3\}$, $\{\boldsymbol{b}_3,\boldsymbol{b}_4\}$ and $\{\boldsymbol{b}_4,\boldsymbol{b}_1\}$ are linearly independent and

$$
\begin{aligned}
C^{(1)} \cap C^{(2)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} \le 0,\ \boldsymbol{b}_2^\top \boldsymbol{y} = 0\},\\
C^{(2)} \cap C^{(3)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_2^\top \boldsymbol{y} \ge 0,\ \boldsymbol{b}_3^\top \boldsymbol{y} = 0\},\\
C^{(3)} \cap C^{(4)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_3^\top \boldsymbol{y} \ge 0,\ \boldsymbol{b}_4^\top \boldsymbol{y} = 0\},\\
C^{(4)} \cap C^{(1)} &= \{\boldsymbol{y} \in \mathbb{R}^{N+1};\ \boldsymbol{b}_1^\top \boldsymbol{y} = 0,\ \boldsymbol{b}_4^\top \boldsymbol{y} \ge 0\}.
\end{aligned}
$$

(Fig. 7).

The scenarios with one and three solution rays lead to contradictions, again, whereas the ones with two or four solution rays are admissible with the same orientation of the rays as in Case IV. In addition, the solution set may contain only the zero vector in this case.

Examples:
(i) Trivial solution set:

$$
\begin{aligned}
F(\boldsymbol{y}) &= -y_1 - y_2 + 2(y_1)_+ + 2(y_2)_+\\
&= \begin{cases} -y_1 - y_2 & \text{if } y_1 \le 0,\ y_2 \le 0,\\ y_1 - y_2 & \text{if } y_1 \ge 0,\ y_2 \le 0,\\ y_1 + y_2 & \text{if } y_1 \ge 0,\ y_2 \ge 0,\\ -y_1 + y_2 & \text{if } y_1 \le 0,\ y_2 \ge 0. \end{cases}
\end{aligned}
$$

(ii) Solution rays in two adjacent cones:

$$
\begin{aligned}
F(\boldsymbol{y}) &= y_1 - y_2 + 2(y_1)_+ + 2(y_2)_+\\
&= \begin{cases} y_1 - y_2 & \text{if } y_1 \le 0,\ y_2 \le 0,\\ 3y_1 - y_2 & \text{if } y_1 \ge 0,\ y_2 \le 0,\\ 3y_1 + y_2 & \text{if } y_1 \ge 0,\ y_2 \ge 0,\\ y_1 + y_2 & \text{if } y_1 \le 0,\ y_2 \ge 0. \end{cases}
\end{aligned}
$$

(ii) Solution rays in two opposite cones:

$$
\begin{aligned}
F(\boldsymbol{y}) &= y_1 + y_2 + 2(y_1)_+ + 2(y_2)_+\\
&= \begin{cases} y_1 + y_2 & \text{if } y_1 \le 0,\ y_2 \le 0,\\ 3y_1 + y_2 & \text{if } y_1 \ge 0,\ y_2 \le 0,\\ 3y_1 + 3y_2 & \text{if } y_1 \ge 0,\ y_2 \ge 0,\\ y_1 + 3y_2 & \text{if } y_1 \le 0,\ y_2 \ge 0. \end{cases}
\end{aligned}
$$

(ii) Four solution rays:

$$
\begin{aligned}
F(\boldsymbol{y}) &= -y_1 + y_2 + 2(y_1)_+ - 2(y_2)_+\\
&= \begin{cases} -y_1 + y_2 & \text{if } y_1 \le 0,\ y_2 \le 0,\\ y_1 + y_2 & \text{if } y_1 \ge 0,\ y_2 \le 0,\\ y_1 - y_2 & \text{if } y_1 \ge 0,\ y_2 \ge 0,\\ -y_1 - y_2 & \text{if } y_1 \le 0,\ y_2 \ge 0. \end{cases}
\end{aligned}
$$

2.3 Bifurcation Criterion

We have observed in the previous subsection that incoherent orientation of solution rays occurs only in the scenarios with *four* branches, that is, only when $\mathbf{0}$ is a bifurcation point of (NF). This leads us to a criterion for detecting border-collision bifurcation points of Problem ($\mathscr{P}$), which we shall describe in this subsection. We shall use the notation from Subsection 2.1.

Let Assumptions 1, 2 and 3 hold and let $\boldsymbol{c}\colon J \to \mathbb{R}^{N+1}$ be a solution curve of ($\mathscr{P}$) defined on an open interval J containing $\bar{s}$ by

$$
\boldsymbol{c}(s) = \begin{cases} \boldsymbol{c}^{(i_1)}(\bar{s} - s) + \bar{\boldsymbol{y}} & \text{if } s < \bar{s},\\ \bar{\boldsymbol{y}} & \text{if } s = \bar{s},\\ \boldsymbol{c}^{(i_2)}(s - \bar{s}) + \bar{\boldsymbol{y}} & \text{if } s > \bar{s} \end{cases}
$$

for some C^1-curves $\boldsymbol{c}^{(i_1)}$ and $\boldsymbol{c}^{(i_2)}$ from Theorem 2.

Plainly,

$$
\boldsymbol{c}'_-(\bar{s}) = -\big(\boldsymbol{c}^{(i_1)}\big)'_+(0), \quad \boldsymbol{c}'_+(\bar{s}) = \big(\boldsymbol{c}^{(i_2)}\big)'_+(0),
$$

and according to Theorem 2,

$$
-\boldsymbol{c}'_-(\bar{s}) \in \operatorname{Ker} \boldsymbol{A}^{(i_1)} \cap \mathring{C}^{(i_1)}, \quad \boldsymbol{c}'_+(\bar{s}) \in \operatorname{Ker} \boldsymbol{A}^{(i_2)} \cap \mathring{C}^{(i_2)}.
$$

By (23) and (7),

$$
\begin{aligned}
\boldsymbol{F}(-\boldsymbol{c}'_-(\bar{s})) &= -\boldsymbol{A}^{(i_1)} \boldsymbol{c}'_-(\bar{s}) = -\boldsymbol{\nabla} \boldsymbol{H}^{(i_1)}(\boldsymbol{c}(\bar{s}))\boldsymbol{c}'_-(\bar{s}) = \mathbf{0},\\
\boldsymbol{F}(\boldsymbol{c}'_+(\bar{s})) &= \boldsymbol{A}^{(i_2)} \boldsymbol{c}'_+(\bar{s}) = \boldsymbol{\nabla} \boldsymbol{H}^{(i_2)}(\boldsymbol{c}(\bar{s}))\boldsymbol{c}'_+(\bar{s}) = \mathbf{0},
\end{aligned}
$$

which particularly means that both $-\boldsymbol{c}'_-(\bar{s})$ and $\boldsymbol{c}'_+(\bar{s})$ solve (NF).

The analysis from the previous subsection suggests us to introduce a **bifurcation criterion** for Problem ($\mathscr{P}$) at the

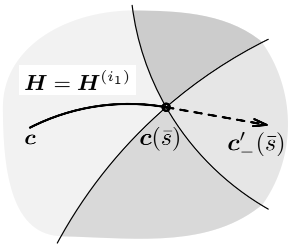


Fig. 8 Determination of the region for seeking a new smooth branch.

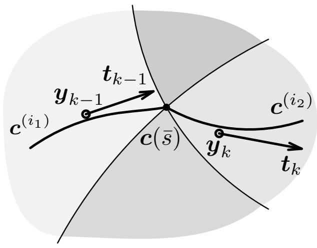


Fig. 9 Numerical continuation of $\boldsymbol{c}$.

border-collision solution $\boldsymbol{c}(\bar{s})$ as the condition on *incoherent* orientation of the one-sided tangents $-\boldsymbol{c}'_-(\bar{s})$ and $\boldsymbol{c}'_+(\bar{s})$:

$$\det\begin{pmatrix}\nabla \boldsymbol{H}^{(i_1)}(\boldsymbol{c}(\bar{s}))\\ \boldsymbol{c}'_-(\bar{s})^\top\end{pmatrix}\det\begin{pmatrix}\nabla \boldsymbol{H}^{(i_2)}(\boldsymbol{c}(\bar{s}))\\ \boldsymbol{c}'_+(\bar{s})^\top\end{pmatrix} < 0. \tag{32}$$

Recall that in the cases corresponding to the simplest scenarios, this condition is never satisfied when $\boldsymbol{c}$ is composed of the only two smooth solution branches of ($\mathscr{P}$) emanating from $\bar{\boldsymbol{y}}$. It is satisfied only if there are four smooth solution branches and $\boldsymbol{c}$ is formed by the branches from mutually opposite regions. In addition, consider a smooth solution curve $\boldsymbol{c}$ with $\boldsymbol{c}(\bar{s}) = \bar{\boldsymbol{y}}$ such that $\boldsymbol{H}$ is smooth there. Then $\boldsymbol{c}'_-(\bar{s}) = \boldsymbol{c}'_+(\bar{s})$, there is only one selection of $\boldsymbol{H}$ in a vicinity of $\boldsymbol{c}(\bar{s})$, that is, $\boldsymbol{H} = \boldsymbol{H}^{(i_1)} = \boldsymbol{H}^{(i_2)}$, and (32) is clearly satisfied neither.

Let us mention that in [21, Subsection 2.2], we proposed a numerical strategy for finding a new smooth solution branch when one smooth branch ending at a boundary solution $\boldsymbol{c}(\bar{s}) = \bar{\boldsymbol{y}}$ is recovered. The strategy consists in seeking a new branch in the region of smoothness lying in the tangential direction of the recovered branch, see Fig. 8. When using that strategy, one could expect that it is the branch in the opposite region that is to be found the most likely if $\bar{\boldsymbol{y}}$ lies in the intersection of four regions. Thus, the potential bifurcation is likely to be detected by the criterion proposed here.

Finally, consider that one traces numerically the solution curve $\boldsymbol{c}$, namely, that one computes a sequence $\{\boldsymbol{y}_k\}$ of points lying approximately on it, and a sequence $\{\boldsymbol{t}_k\}$ of approximations of the corresponding tangent vectors. We shall suppose that $\boldsymbol{H}$ is smooth at each $\boldsymbol{y}_k$, for simplicity, as it is hardly possible to encounter exactly a point where $\boldsymbol{H}$ is not smooth. Further, let a boundary solution $\boldsymbol{c}(\bar{s})$ lie between $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ such that $\boldsymbol{y}_{k-1}$, $\boldsymbol{y}_k$ approximate some solutions on $\boldsymbol{c}^{(i_1)}$ and $\boldsymbol{c}^{(i_2)}$, respectively (Fig. 9). Assuming that both $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ are close to $\boldsymbol{c}(\bar{s})$ and $\boldsymbol{t}_{k-1}$, $\boldsymbol{t}_k$ are good approximations of $\boldsymbol{c}'_-(\bar{s})$ and $\boldsymbol{c}'_+(\bar{s})$, respectively, we arrive at the following test for detecting border-collision bifurcations based on (32):

$$\det\begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1})\\ \boldsymbol{t}_{k-1}^\top\end{pmatrix}\det\begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_k)\\ \boldsymbol{t}_k^\top\end{pmatrix} < 0. \tag{33}$$

Let us mention that the obtained test is the same as the standard one for detecting the so-called (smooth) simple bifurcation points (see, for example, [4, Section 24]).

3 Numerical Realisation of the Bifurcation Test

The finite-dimensional problem ($\mathscr{P}$) arises typically from discretisation of a parameter-dependent problem of infinite dimension and its size N may become quite large. In such cases, it may not be clear how to determine numerically the sign of the product of the determinants appearing in (33). The aim of this section is to propose a technique requiring resolution of a sequence of linear systems instead.

Our approach is based on the following observation, which was employed for a test of simple bifurcation points in [15].

Lemma 2 *Let $\boldsymbol{J} \in \mathbb{R}^{(N+1)\times(N+1)}$, $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbb{R}^{N+1}$ and $d \in \mathbb{R}$ be such that the matrix $\boldsymbol{M} \in \mathbb{R}^{(N+2)\times(N+2)}$ defined by*

$$\boldsymbol{M} := \begin{pmatrix}\boldsymbol{J} & \boldsymbol{b}\\ \boldsymbol{c}^\top & d\end{pmatrix} \tag{34}$$

is non-singular. Introduce $\tau \in \mathbb{R}$ implicitly via

$$\boldsymbol{M}\begin{pmatrix}\boldsymbol{v}\\ \tau\end{pmatrix} = \begin{pmatrix}\boldsymbol{0}\\ 1\end{pmatrix}.$$

Then

$$\tau = \frac{\det \boldsymbol{J}}{\det \boldsymbol{M}}.$$

Proof The assertion follows directly from Cramer's rule.

□

Next, let $\boldsymbol{y}_{k-1}$, $\boldsymbol{y}_k$, $\boldsymbol{t}_{k-1}$ and $\boldsymbol{t}_k$ be given so that both determinants appearing in (33) are non-zero, and let $\boldsymbol{b}$, $\boldsymbol{c} \in \mathbb{R}^{N+1}$ and $d \in \mathbb{R}$ be fixed. Set

$$\boldsymbol{J}(\alpha) := (1-\alpha)\begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1})\\ \boldsymbol{t}_{k-1}^\top\end{pmatrix} + \alpha\begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_k)\\ \boldsymbol{t}_k^\top\end{pmatrix}, \tag{35}$$

$$\boldsymbol{M}(\alpha) := \begin{pmatrix}\boldsymbol{J}(\alpha) & \boldsymbol{b}\\ \boldsymbol{c}^\top & d\end{pmatrix} \tag{36}$$

for any $\alpha \in [0,1]$, and whenever $\boldsymbol{M}(\alpha)$ is non-singular, take $\tau(\alpha)$ satisfying

$$\boldsymbol{M}(\alpha)\begin{pmatrix}\boldsymbol{v}(\alpha)\\ \tau(\alpha)\end{pmatrix} = \begin{pmatrix}\boldsymbol{0}\\ 1\end{pmatrix}. \tag{37}$$

Then

$$\tau(\alpha) = \frac{\det \boldsymbol{J}(\alpha)}{\det \boldsymbol{M}(\alpha)},$$

and observing that

$$\boldsymbol{J}(0) = \begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1})\\ \boldsymbol{t}_{k-1}^\top\end{pmatrix}, \quad \boldsymbol{J}(1) = \begin{pmatrix}\nabla \boldsymbol{H}(\boldsymbol{y}_k)\\ \boldsymbol{t}_k^\top\end{pmatrix},$$

one can test (33) equivalently by comparing the signs of $\det \boldsymbol{J}(0)$ and $\det \boldsymbol{J}(1)$.

Suppose for a while that $\boldsymbol{b}$, $\boldsymbol{c}$ and d are chosen so that

$$\det \boldsymbol{J}(\alpha) = 0 \implies \det \boldsymbol{M}(\alpha) \neq 0, \quad \alpha \in (0,1), \tag{38}$$

$$\det \boldsymbol{M}(0), \det \boldsymbol{M}(1) \neq 0. \tag{39}$$

Then $\det \boldsymbol{M}(\alpha)$ is a non-zero polynomial in α, and $\tau(\alpha)$ is a well-defined function with $\tau(0)$ and $\tau(1)$ finite. Furthermore, the sign changes of $\det \boldsymbol{J}(\alpha)$ are characterised by passing of $\tau(\alpha)$ through 0 whereas the sign changes of $\det \boldsymbol{M}(\alpha)$ by sign changes of $\tau(\alpha)$ caused by a singularity. Since these two cases can be easily distinguished, we are lead to the following idea for testing (33): To follow the behaviour of $\tau(\alpha)$ when α passes through $[0,1]$ and to monitor the sign changes of $\det \boldsymbol{J}(\alpha)$.

It remains to find suitable $\boldsymbol{b}$, $\boldsymbol{c}$ and d. Our choice will be based on a classical result [13], [24, Lemma 5.6]:

Lemma 3 *Let $\boldsymbol{M} \in \mathbb{R}^{(N+2)\times(N+2)}$ be given by* (34)*. Necessary and sufficient conditions for its non-singularity are:*

(i) $d \neq \boldsymbol{c}^\top \boldsymbol{J}^{-1} \boldsymbol{b}$ *if $\boldsymbol{J}$ is non-singular.*

(ii) $\dim \operatorname{Ker} \boldsymbol{J} = 1, \boldsymbol{b} \notin \operatorname{Im} \boldsymbol{J}$ *and* $\boldsymbol{c} \notin \operatorname{Im} \boldsymbol{J}^\top$ *if $\boldsymbol{J}$ is singular.*

As $\det \boldsymbol{J}(0)$ and $\det \boldsymbol{J}(1)$ are supposed to be non-zero, $\det \boldsymbol{J}(\alpha)$ is a non-zero polynomial in α of order at most $N+1$, and it has a finite number of roots in $(0,1)$ (possibly zero). Denoting them $\alpha_1, \ldots, \alpha_{n_r}$, one can ensure (38) and (39) by the following assumption:

Assumption 4 *Let*

(i) $\dim \ker \boldsymbol{J}(\alpha_i) = 1, \quad i = 1, \ldots, n_r$;

(ii) $\boldsymbol{b} \notin \operatorname{Im} \boldsymbol{J}(\alpha_i), \boldsymbol{c} \notin \operatorname{Im} \boldsymbol{J}^\top(\alpha_i), \quad i = 1, \ldots, n_r$;

(iii) $d \neq \boldsymbol{c}^\top \boldsymbol{J}^{-1}(0) \boldsymbol{b}, \boldsymbol{c}^\top \boldsymbol{J}^{-1}(1) \boldsymbol{b}$.

To start with investigation of these conditions, we shall analyse the part (i) in the cases of the simplest scenarios from Subsection 2.2.

Firstly, consider that $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ in question belong to solution branches in adjacent regions of $\boldsymbol{H}$ meeting at $\bar{\boldsymbol{y}}$ and corresponding to one of Cases I–V. Without loss of generality, one can write

$$\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1}) = \nabla \boldsymbol{H}^{(1)}(\boldsymbol{y}_{k-1}), \quad \nabla \boldsymbol{H}(\boldsymbol{y}_k) = \nabla \boldsymbol{H}^{(2)}(\boldsymbol{y}_k),$$

and setting

$$\boldsymbol{A}^{(1)} = \nabla \boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}}), \quad \boldsymbol{A}^{(2)} = \nabla \boldsymbol{H}^{(2)}(\bar{\boldsymbol{y}}),$$

one has according to the analysis in Subsection 2.2 that

$$\operatorname{rank} \boldsymbol{A}^{(1)} = \operatorname{rank} \boldsymbol{A}^{(2)} = N,$$

$$\boldsymbol{A}^{(1)} \boldsymbol{y} = \boldsymbol{A}^{(2)} \boldsymbol{y},$$

$$\forall \boldsymbol{y} \in \mathbb{R}^{N+1} \text{ with } \boldsymbol{b}^\top \boldsymbol{y} = 0 \text{ for some } \boldsymbol{0} \neq \boldsymbol{b} \in \mathbb{R}^{N+1},$$

$$\begin{pmatrix} \boldsymbol{A}^{(1)} \\ \boldsymbol{b}^\top \end{pmatrix} \text{ is non-singular.}$$

Proceeding as in the proof of Lemma 1 and introducing an orthonormal basis $\{\boldsymbol{q}_1, \ldots \boldsymbol{q}_{N+1}\}$ of $\mathbb{R}^{N+1}$ with $\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, one can define

$$\hat{\boldsymbol{A}}^{(i)} := \boldsymbol{A}^{(i)} \boldsymbol{Q}^\top, \quad i = 1,2,$$

with the rows of $\boldsymbol{Q} \in \mathbb{R}^{(N+1)\times(N+1)}$ formed by $\boldsymbol{q}_1^\top, \ldots, \boldsymbol{q}_{N+1}^\top$. One deduces as before that

$$\hat{\boldsymbol{A}}^{(i)} = (\hat{\boldsymbol{A}}_1^{(i)}, \hat{\boldsymbol{A}}_2), \quad i = 1,2,$$

where $\hat{\boldsymbol{A}}_1^{(i)} \in \mathbb{R}^N$, $\hat{\boldsymbol{A}}_2 \in \mathbb{R}^{N\times N}$ and $\operatorname{rank} \hat{\boldsymbol{A}}_2 = N$. For an arbitrary $\alpha \in (0,1)$,

$$\begin{aligned}\big((1-\alpha)\hat{\boldsymbol{A}}_1^{(1)} + \alpha \hat{\boldsymbol{A}}_1^{(2)}, \hat{\boldsymbol{A}}_2\big) &= (1-\alpha)\hat{\boldsymbol{A}}^{(1)} + \alpha \hat{\boldsymbol{A}}^{(2)} \\ &= \big((1-\alpha)\boldsymbol{A}^{(1)} + \alpha \boldsymbol{A}^{(2)}\big)\boldsymbol{Q}^\top\end{aligned}$$

and consequently,

$$\begin{aligned} N &= \operatorname{rank}\big((1-\alpha)\hat{\boldsymbol{A}}^{(1)} + \alpha \hat{\boldsymbol{A}}^{(2)}\big) \\ &\leq \min\big\{\operatorname{rank}\big((1-\alpha)\boldsymbol{A}^{(1)} + \alpha \boldsymbol{A}^{(2)}\big), \operatorname{rank} \boldsymbol{Q}^\top\big\}, \\ N &\leq \operatorname{rank}\big((1-\alpha)\nabla \boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}}) + \alpha \nabla \boldsymbol{H}^{(2)}(\bar{\boldsymbol{y}})\big).\end{aligned}$$

If both $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ are sufficiently close to $\bar{\boldsymbol{y}}$, standard continuity argumentation yields that

$$N \leq \operatorname{rank} \begin{pmatrix} (1-\alpha_i)\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1}) + \alpha_i \nabla \boldsymbol{H}(\boldsymbol{y}_k) \\ (1-\alpha_i)\boldsymbol{t}_{k-1}^\top + \alpha_i \boldsymbol{t}_k^\top \end{pmatrix} = \operatorname{rank} \boldsymbol{J}(\alpha_i), \quad i = 1, \ldots, n_r.$$

This shows that Assumption 4(i) is always satisfied for $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ from adjacent regions.

Secondly, consider that $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ belong to solution branches in opposite regions meeting at $\bar{\boldsymbol{y}}$ and corresponding to Case IV or V. Let

$$\nabla \boldsymbol{H}(\boldsymbol{y}_{k-1}) = \nabla \boldsymbol{H}^{(1)}(\boldsymbol{y}_{k-1}), \quad \nabla \boldsymbol{H}(\boldsymbol{y}_k) = \nabla \boldsymbol{H}^{(3)}(\boldsymbol{y}_k),$$

$$\boldsymbol{A}^{(i)} = \nabla \boldsymbol{H}^{(i)}(\bar{\boldsymbol{y}}), \quad i = 1,2,3.$$

Similarly as before, one can suppose that

$$\operatorname{rank} \boldsymbol{A}^{(i)} = N, \quad i = 1,3,$$

and find linearly independent $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathbb{R}^{N+1}$ such that

$$\boldsymbol{A}^{(1)} \boldsymbol{y} = \boldsymbol{A}^{(2)} \boldsymbol{y}, \ \forall \boldsymbol{y} \in \mathbb{R}^{N+1} \text{ with } \boldsymbol{b}_1^\top \boldsymbol{y} = 0,$$

$$\boldsymbol{A}^{(2)} \boldsymbol{y} = \boldsymbol{A}^{(3)} \boldsymbol{y}, \ \forall \boldsymbol{y} \in \mathbb{R}^{N+1} \text{ with } \boldsymbol{b}_2^\top \boldsymbol{y} = 0,$$

$$\begin{pmatrix} \boldsymbol{A}^{(1)} \\ \boldsymbol{b}_1^\top \end{pmatrix} \text{ is non-singular.}$$

In this case, one can take $\boldsymbol{q}_1 := \boldsymbol{b}_1/\|\boldsymbol{b}_1\|$, $\boldsymbol{q}_2 := \boldsymbol{b}_2/\|\boldsymbol{b}_2\|$, pick out an orthonormal basis $\{\boldsymbol{q}_3, \ldots \boldsymbol{q}_{N+1}\}$ of $\{\boldsymbol{q}_1, \boldsymbol{q}_2\}^\perp$, and compose $\boldsymbol{Q} \in \mathbb{R}^{(N+1)\times(N+1)}$ from the row vectors $\boldsymbol{q}_1^\top, \ldots, \boldsymbol{q}_{N+1}^\top$. Then one can define

$$\hat{\boldsymbol{A}}^{(i)} := \boldsymbol{A}^{(i)}\boldsymbol{Q}^{-1}, \quad i = 1,2,3. \tag{40}$$

It is easy to verify that the columns of $\boldsymbol{Q}^{-1}$ are $\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2, \boldsymbol{q}_3, \ldots, \boldsymbol{q}_{N+1}$ with $\tilde{\boldsymbol{q}}_1$ and $\tilde{\boldsymbol{q}}_2$ determined uniquely by

$$\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2 \in \operatorname{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2\}, \quad \boldsymbol{q}_i^\top \tilde{\boldsymbol{q}}_j = \delta_{ij},\ i,j = 1,2.$$

Further, one can show that

$$\hat{\boldsymbol{A}}^{(1)} = (\hat{\boldsymbol{A}}_1^{(1)}, \hat{\boldsymbol{A}}_2^{(1)}, \hat{\boldsymbol{A}}_3), \quad \hat{\boldsymbol{A}}^{(2)} = (\hat{\boldsymbol{A}}_1^{(3)}, \hat{\boldsymbol{A}}_2^{(1)}, \hat{\boldsymbol{A}}_3),$$
$$\hat{\boldsymbol{A}}^{(3)} = (\hat{\boldsymbol{A}}_1^{(3)}, \hat{\boldsymbol{A}}_2^{(3)}, \hat{\boldsymbol{A}}_3)$$

for some $\hat{\boldsymbol{A}}_1^{(1)}, \hat{\boldsymbol{A}}_2^{(1)}, \hat{\boldsymbol{A}}_1^{(3)}, \hat{\boldsymbol{A}}_2^{(3)} \in \mathbb{R}^N$, $\hat{\boldsymbol{A}}_3 \in \mathbb{R}^{N\times(N-1)}$ and $\operatorname{rank}\hat{\boldsymbol{A}}_3 = N-1$. As a consequence, if both $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ are sufficiently close to $\bar{\boldsymbol{y}}$,

$$N-1 \le \operatorname{rank}\begin{pmatrix} (1-\alpha_i)\nabla\boldsymbol{H}(\boldsymbol{y}_{k-1}) + \alpha_i\nabla\boldsymbol{H}(\boldsymbol{y}_k) \\ (1-\alpha_i)\boldsymbol{t}_{k-1}^\top + \alpha_i\boldsymbol{t}_k^\top \end{pmatrix} = \operatorname{rank}\boldsymbol{J}(\alpha_i), \quad i = 1,\ldots,n_r,$$

that is, $\dim\ker\boldsymbol{J}(\alpha_i) \le 2$. Assumption 4(i) may be violated for some α_i, in general.

Nevertheless, if $\dim\ker\boldsymbol{J}(\alpha) = 2$ for some $\alpha \in (0,1)$, then

$$(1-\alpha)\hat{\boldsymbol{A}}_j^{(1)} + \alpha\hat{\boldsymbol{A}}_j^{(3)} \in \operatorname{Im}\hat{\boldsymbol{A}}_3,$$
$$\hat{\boldsymbol{A}}_j^{(3)} \in \frac{\alpha-1}{\alpha}\hat{\boldsymbol{A}}_j^{(1)} + \operatorname{Im}\hat{\boldsymbol{A}}_3$$

for $j = 1,2$. By (40),

$$\hat{\boldsymbol{A}}_j^{(i)} = \boldsymbol{A}^{(i)}\tilde{\boldsymbol{q}}_j = \nabla\boldsymbol{H}^{(i)}(\bar{\boldsymbol{y}})\tilde{\boldsymbol{q}}_j, \quad i = 1,3,\ j = 1,2,$$
$$\hat{\boldsymbol{A}}_3 = \boldsymbol{A}^{(1)}\boldsymbol{Q}_3^{-1} = \nabla\boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}})\boldsymbol{Q}_3^{-1},$$

where the columns of $\boldsymbol{Q}_3^{-1} \in \mathbb{R}^{(N+1)\times(N-1)}$ are formed by $\boldsymbol{q}_3, \ldots, \boldsymbol{q}_{N+1}$. From here, a necessary condition for violation of Assumption 4(i) reads

$$\nabla\boldsymbol{H}^{(3)}(\bar{\boldsymbol{y}})\tilde{\boldsymbol{q}}_j \in r\nabla\boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}})\tilde{\boldsymbol{q}}_j + \operatorname{Im}\big(\nabla\boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}})\boldsymbol{Q}_3^{-1}\big)$$

for some $r \in (-\infty, 0)$, r the *same* for both $j = 1,2$. One can conclude that this situation seems to be rare: if it satisfied for $j = 1$ and some r, $\nabla\boldsymbol{H}^{(3)}(\bar{\boldsymbol{y}})\tilde{\boldsymbol{q}}_2$ has to lie in $r\nabla\boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}})\tilde{\boldsymbol{q}}_2 + \operatorname{Im}\big(\nabla\boldsymbol{H}^{(1)}(\bar{\boldsymbol{y}})\boldsymbol{Q}_3^{-1}\big)$, which is an $(N-1)$-dimensional affine space in $\mathbb{R}^N$.

Once (i) in Assumption 4 being fulfilled, both $\operatorname{Im}\boldsymbol{J}(\alpha_i)$ and $\operatorname{Im}\boldsymbol{J}^\top(\alpha_i)$ from (ii) are N-dimensional subspaces of $\mathbb{R}^{N+1}$ for any $i = 1,\ldots,n_r$. Therefore, (ii) is satisfied whenever $\boldsymbol{b}$ and $\boldsymbol{c}$ are out of certain finite unions of N-dimensional subspaces of $\mathbb{R}^{N+1}$.

Finally, if $\boldsymbol{b}$ and $\boldsymbol{c}$ are chosen in accordance with (ii), (iii) is met for d different from two specific values.

One can conclude that there seem to be practically no restrictions on the choice of $\boldsymbol{b}$, $\boldsymbol{c}$ and d so one can suppose Assumption 4 to hold when choosing them randomly. We propose the following numerical procedure for testing Condition (33):

Algorithm 1

Input: $\boldsymbol{y}_{k-1}, \boldsymbol{y}_k, \boldsymbol{t}_{k-1}, \boldsymbol{t}_k \in \mathbb{R}^{N+1}$, $\delta_{\max} > \delta_{\min} > 0$, $\delta_{\text{inc}} > 1 > \delta_{\text{dec}} > 0$.

Step 1: Set $n_{\text{ch}} := 0$, $\alpha := 0$, $\tau_0 := 10^6$, $\tau_{-1} := 10^6$, $\delta := \delta_{\min}$, and choose $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^{N+1}$ and $d \in \mathbb{R}$ randomly.

Step 2: Set

$$\boldsymbol{J} := (1-\alpha)\begin{pmatrix} \nabla\boldsymbol{H}(\boldsymbol{y}_{k-1}) \\ \boldsymbol{t}_{k-1}^\top \end{pmatrix} + \alpha\begin{pmatrix} \nabla\boldsymbol{H}(\boldsymbol{y}_k) \\ \boldsymbol{t}_k^\top \end{pmatrix}, \quad \boldsymbol{M} := \begin{pmatrix} \boldsymbol{J} & \boldsymbol{b} \\ \boldsymbol{c}^\top & d \end{pmatrix}.$$

Step 3: Solve

$$\boldsymbol{M}\begin{pmatrix} \boldsymbol{v} \\ \tau \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ 1 \end{pmatrix}.$$

Step 4: If $\tau\tau_0 < 0$ and $|\tau_0| < |\tau_{-1}|$, set $n_{\text{ch}} := n_{\text{ch}} + 1$.

Step 5: If $|\tau - \tau_0|$ is large, set $\delta := \max\{\delta_{\text{dec}}\delta, \delta_{\min}\}$, otherwise if $|\tau - \tau_0|$ is small, set $\delta := \min\{\delta_{\text{inc}}\delta, \delta_{\max}\}$.

Step 6: If $\alpha < 1$, set $\alpha := \min\{\alpha + \delta, 1\}$, $\tau_{-1} := \tau_0$, $\tau_0 := \tau$ and go to Step 2.

Step 7: If n_{ch} is odd, print "bifurcation detected".

Here, the last two values of τ are stored in τ_0 and τ_{-1}, and n_{ch} serves for counting the sign changes of $\det\boldsymbol{J}$. It is increased only if $\tau\tau_0 < 0$ and $|\tau_0| < |\tau_{-1}|$ simultaneously because it is expected that τ has passed through a singularity in the case of $|\tau_0| \ge |\tau_{-1}|$. Therefore, odd n_{ch} at the end indicates opposite signs of $\det\begin{pmatrix} \nabla\boldsymbol{H}(\boldsymbol{y}_{k-1}) \\ \boldsymbol{t}_{k-1}^\top \end{pmatrix}$ and $\det\begin{pmatrix} \nabla\boldsymbol{H}(\boldsymbol{y}_k) \\ \boldsymbol{t}_k^\top \end{pmatrix}$.

The values of α are increased by the current values of the variable δ. The latter one is bounded by $\delta_{\min}$ and $\delta_{\max}$ and adapted by the scale factors δ_{inc} and δ_{dec} according to the latest values of τ. This adaptivity serves for effective treatment of the singularities of τ, which are characterised by large $|\tau - \tau_0|$. Since it is highly improbable to encounter exactly a value of α with $\boldsymbol{M}(\alpha)$ singular, we take no care of this possibility.

4 Application to Plane Contact Problems with Friction

Let us consider static deformation of an elastic body whose reference configuration is the closure of a bounded domain $\Omega \subset \mathbb{R}^2$. Let the boundary $\partial\Omega$ be Lipschitz-continuous and split into three disjoint relatively open subsets Γ_D, Γ_N and Γ_c. The deformation of the body will be denoted by $\boldsymbol{\varphi}$, and it can be expressed alternatively in terms of the displacement $\boldsymbol{u}$ as $\boldsymbol{\varphi} = \boldsymbol{id} + \boldsymbol{u}$, where $\boldsymbol{id}$ stands for the identity mapping.

The displacement $\boldsymbol{u}_D$ being imposed on Γ_D, the body in the deformed configuration $\boldsymbol{\varphi}(\Omega)$ is subject to body forces of the density $\boldsymbol{f}^{\boldsymbol{\varphi}}$, and surface forces of the density $\boldsymbol{h}^{\boldsymbol{\varphi}}$ act on

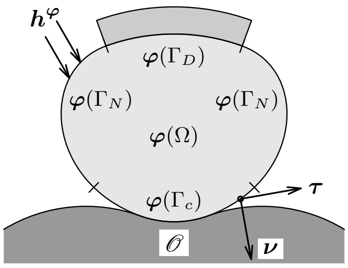


Fig. 10 Geometry of the problem.

$\boldsymbol{\varphi}(\Gamma_N)$. We suppose that the prescribed displacement and the forces depend on a real parameter γ, that is, $\boldsymbol{u}_D = \boldsymbol{u}_D(\boldsymbol{x},\gamma)$, $\boldsymbol{f}^{\boldsymbol{\varphi}} = \boldsymbol{f}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}},\gamma)$ and $\boldsymbol{h}^{\boldsymbol{\varphi}} = \boldsymbol{h}^{\boldsymbol{\varphi}}(\boldsymbol{x}^{\boldsymbol{\varphi}},\gamma)$, where $\boldsymbol{x}$ and $\boldsymbol{x}^{\boldsymbol{\varphi}}$ stand for points in the initial and the deformed configuration, respectively.

The points from Γ_c may come into contact with a fixed curved rigid obstacle represented by a closed set $\mathscr{O} \subset \mathbb{R}^2$, the contact being described by unilateral conditions and the Coulomb friction law. We assume that there exist a neighbourhood V of $\partial\mathscr{O}$ and a differentiable function $g\colon V \to \mathbb{R}$ such that $\boldsymbol{\varphi}(\Gamma_c) \subset V$, $\nabla g \neq \boldsymbol{0}$ in V and

$$\begin{aligned} &g(\boldsymbol{x}) > 0, \quad \forall \boldsymbol{x} \in V \cap \operatorname{int}\mathscr{O},\\ &g(\boldsymbol{x}) = 0, \quad \forall \boldsymbol{x} \in \partial\mathscr{O},\\ &g(\boldsymbol{x}) < 0, \quad \forall \boldsymbol{x} \in V \cap \operatorname{ext}\mathscr{O}. \end{aligned}$$

Then one can extend the unit inward normal $\boldsymbol{\nu}$ and the unit tangent $\boldsymbol{\tau}$ to the obstacle from $\partial\mathscr{O}$ to V as

$$\boldsymbol{\nu}(\boldsymbol{x}) = \frac{\nabla g(\boldsymbol{x})}{\|\nabla g(\boldsymbol{x})\|}, \quad \boldsymbol{\tau}(\boldsymbol{x}) = (-\nu_2(\boldsymbol{x}), \nu_1(\boldsymbol{x}))$$

(see Fig. 10).

The classical formulation of the parametrised equilibrium problem reads as follows:

$$\left.\begin{aligned} &\textit{Find } \boldsymbol{u} \in \mathscr{U}_{\text{ad}} \textit{ such that}\\ &\operatorname{\mathbf{div}} \boldsymbol{\sigma}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x},\gamma) = \boldsymbol{0},\ \boldsymbol{x} \in \Omega,\\ &\boldsymbol{\sigma}(\boldsymbol{x}) = \hat{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{I} + \nabla\boldsymbol{u}(\boldsymbol{x})),\ \boldsymbol{x} \in \Omega,\\ &\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{u}_D(\boldsymbol{x},\gamma),\ \boldsymbol{x} \in \Gamma_D,\\ &\boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x},\gamma),\ \boldsymbol{x} \in \Gamma_N,\\ &\left.\begin{aligned} g(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})) \le 0,\ T_\nu(\boldsymbol{x}) \le 0,\\ g(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))T_\nu(\boldsymbol{x}) = 0, \end{aligned}\right\} \boldsymbol{x} \in \Gamma_c,\\ &\left.\begin{aligned} |T_\tau(\boldsymbol{x})| \le -\mathscr{F}(\boldsymbol{x})T_\nu(\boldsymbol{x}),\\ u_\tau(\boldsymbol{x}) \neq 0 \implies T_\tau(\boldsymbol{x}) = \mathscr{F}(\boldsymbol{x})T_\nu(\boldsymbol{x})\frac{u_\tau(\boldsymbol{x})}{|u_\tau(\boldsymbol{x})|}, \end{aligned}\right\} \boldsymbol{x} \in \Gamma_c, \end{aligned}\right\} \quad (41)$$

where γ varies over an interval of interest. The set $\mathscr{U}_{\text{ad}}$ of kinematically admissible displacements is introduced by

$$\begin{aligned} \mathscr{U}_{\text{ad}} := \{\boldsymbol{v}\colon \overline{\Omega} \to \mathbb{R}^2 \text{ “smooth enough”};\\ \boldsymbol{id} + \boldsymbol{v} \text{ is injective in } \Omega,\ \det(\boldsymbol{I} + \nabla\boldsymbol{v}) > 0 \text{ in } \overline{\Omega}\}, \end{aligned}$$

the first Piola-Kirchhoff stress tensor is denoted by $\boldsymbol{\sigma}$, its response function characterising the material of the elastic body by $\hat{\boldsymbol{\sigma}}$ and $\boldsymbol{I}$ is the identity matrix. Further, $\boldsymbol{n}$ stands for the unit outward normal vector along $\partial\Omega$ and

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{x},\gamma) &= \det(\boldsymbol{I} + \nabla\boldsymbol{u}(\boldsymbol{x}))\boldsymbol{f}^{\boldsymbol{\varphi}}(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}),\gamma),\\ \boldsymbol{h}(\boldsymbol{x},\gamma) &= \det(\boldsymbol{I} + \nabla\boldsymbol{u}(\boldsymbol{x}))\|(\boldsymbol{I} + \nabla\boldsymbol{u}(\boldsymbol{x}))^{-\top}\boldsymbol{n}(\boldsymbol{x})\|\boldsymbol{h}^{\boldsymbol{\varphi}}(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}),\gamma) \end{aligned}$$

are the densities of the volume and surface forces related to the reference configuration. Finally,

$$T_\nu(\boldsymbol{x}) = \boldsymbol{T}(\boldsymbol{x}) \cdot \boldsymbol{\nu}(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})), \quad T_\tau(\boldsymbol{x}) = \boldsymbol{T}(\boldsymbol{x}) \cdot \boldsymbol{\tau}(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))$$

are the components of the first Piola-Kirchhoff stress vector $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x})$ in the directions $\boldsymbol{\nu}$ and $\boldsymbol{\tau}$, $\mathscr{F}$ is a non-negative function representing the friction coefficient and

$$u_\tau(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}) \cdot \boldsymbol{\tau}(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))$$

is the tangential displacement.

Discretisation of this problem is done by applying a Lagrange finite-element method to a mixed variational formulation of (41) with Lagrange multipliers enforcing the Dirichlet and the contact boundary conditions. In particular, we consider the nodal approximation of the contact conditions written in terms of projections proposed in [1]. This leads to the discrete problem in the form of $(\mathscr{P})$:

$$\left.\begin{aligned} &\textit{Find } \boldsymbol{y} := (\gamma, \boldsymbol{u}, \boldsymbol{\lambda}_D, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_\tau) \in \mathbb{R}^{1+2(n_\Omega+n_D+n_c)}\\ &\textit{such that}\\ &\boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{0}, \end{aligned}\right\} \quad (42)$$

where $\boldsymbol{H}\colon \mathbb{R}^{1+2(n_\Omega+n_D+n_c)} \to \mathbb{R}^{2(n_\Omega+n_D+n_c)}$ is defined by

$$\boldsymbol{H}(\boldsymbol{y}) := \begin{pmatrix} \boldsymbol{A}(\boldsymbol{u}) - \boldsymbol{L}(\gamma,\boldsymbol{u}) - \boldsymbol{B}_D^\top\boldsymbol{\lambda}_D - \boldsymbol{B}_\nu^\top(\boldsymbol{u})\boldsymbol{\lambda}_\nu - \boldsymbol{B}_\tau^\top(\boldsymbol{u})\boldsymbol{\lambda}_\tau\\ \boldsymbol{B}_D\boldsymbol{u} - \boldsymbol{U}_D(\gamma)\\ \lambda_{\nu,j} - (\lambda_{\nu,j} - g_j(\boldsymbol{u}))_-, \quad j = 1,\dots,n_c\\ \lambda_{\tau,j} - P_{[\mathscr{F}_j(\lambda_{\nu,j}-g_j(\boldsymbol{u}))_-, -\mathscr{F}_j(\lambda_{\nu,j}-g_j(\boldsymbol{u}))_-]}\big(\lambda_{\tau,j}\\ -(\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j\big), \quad j = 1,\dots,n_c \end{pmatrix}.$$

Here $\boldsymbol{u} \in \mathbb{R}^{2n_\Omega}$ is the vector of the nodal displacements, $\boldsymbol{\lambda}_D \in \mathbb{R}^{2n_D}$ is the Lagrange multiplier corresponding to the Dirichlet condition, and $\boldsymbol{\lambda}_\nu \in \mathbb{R}^{n_c}$ and $\boldsymbol{\lambda}_\tau \in \mathbb{R}^{n_c}$ are the normal and tangential Lagrange multipliers on the contact zone, respectively. Furthermore, $\boldsymbol{A}(\boldsymbol{u})$ and $\boldsymbol{L}(\gamma,\boldsymbol{u})$ are the vectors of the internal elastic and external applied forces, respectively, and $\boldsymbol{B}_D$ is the kinematic transformation matrix linking $\boldsymbol{u}$ with the Lagrange multiplier $\boldsymbol{\lambda}_D$. The matrices $\boldsymbol{B}_\nu(\boldsymbol{u})$ and $\boldsymbol{B}_\tau(\boldsymbol{u})$, depending on the actual position of the body, associate the vector of the nodal displacements with the vectors of the nodal displacements in the directions $\boldsymbol{\nu}$ and $\boldsymbol{\tau}$, respectively. The vector $\boldsymbol{U}_D(\gamma)$ corresponds to the prescribed displacement on Γ_D, $g_j(\boldsymbol{u})$ are the values of g for the actual positions of the contact nodes, and $\mathscr{F}_j$ are the values of the friction coefficient in these nodes. Finally, $(.)_-$ and $P_{[a,b]}(.)$ stand for the projections onto the intervals $(-\infty,0]$ and $[a,b]$, respectively.

Since both projections involved are PC^1-function, $\boldsymbol{H}$ is also PC^1 under the following assumption:

Assumption 5 *Let $\boldsymbol{A}$, $\boldsymbol{L}$ and $\boldsymbol{U}_D$ be C^1-functions and g be of class C^2.*

Then one can construct selections $\boldsymbol{H}^{(i)}$ of $\boldsymbol{H}$ following [5], the idea being similar to finding possible evolutions of the quasi-static problem in [11]:

Let $\bar{\boldsymbol{y}} = (\bar{\gamma}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{\lambda}}_D, \bar{\boldsymbol{\lambda}}_\nu, \bar{\boldsymbol{\lambda}}_\tau) \in \mathbb{R}^{1+2(n_\Omega+n_D+n_c)}$ be an arbitrary but fixed point and introduce subsets of the index set of all contact nodes $I := \{1,\dots,n_c\}$ as follows:

$$
\begin{aligned}
I_\mathrm{f} &= \{j \in I;\ \bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}) > 0\},\\
I_\mathrm{c} &= \{j \in I;\ \bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}) < 0\},\\
I_\mathrm{z} &= \{j \in I;\ \bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}) = 0\},\\
I_0 &= \{j \in I;\ \mathscr{F}_j = 0\},\\
I_\mathrm{s} &= \{j \in I;\\
&\qquad \mathscr{F}_j > 0,\ |\bar{\lambda}_{\tau,j} - (\boldsymbol{B}_\tau(\bar{\boldsymbol{u}})\bar{\boldsymbol{u}})_j| < -\mathscr{F}_j(\bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}))_-\},\\
I_\mathrm{l} &= \{j \in I;\\
&\qquad \mathscr{F}_j > 0,\ |\bar{\lambda}_{\tau,j} - (\boldsymbol{B}_\tau(\bar{\boldsymbol{u}})\bar{\boldsymbol{u}})_j| > -\mathscr{F}_j(\bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}))_-\},\\
I_\mathrm{i} &= \{j \in I;\\
&\qquad \mathscr{F}_j > 0,\ |\bar{\lambda}_{\tau,j} - (\boldsymbol{B}_\tau(\bar{\boldsymbol{u}})\bar{\boldsymbol{u}})_j| = -\mathscr{F}_j(\bar{\lambda}_{\nu,j} - g_j(\bar{\boldsymbol{u}}))_-\}.
\end{aligned}
$$

Apparently, if $\bar{\boldsymbol{y}}$ is a solution of (42), I_f is composed of the indices of the nodes not in contact (free), I_c of the indices of the nodes in strong contact, I_z of the indices of the nodes in grazing contact (with zero contact forces), I_0 of the indices of the nodes with vanishing friction, I_s of the indices of the nodes in strong stick, I_l of the indices of the nodes in non-zero slip, and I_i of the indices of the nodes in impending slip.

In virtue of Assumption 5, the functions $\boldsymbol{u} \mapsto g_j(\boldsymbol{u})$ and $\boldsymbol{u} \mapsto \boldsymbol{B}_\tau(\boldsymbol{u})$ are continuous. It is then readily seen that if $I_\mathrm{z} \cup (I_\mathrm{c} \cap I_\mathrm{i}) = \emptyset$, there exists a neighbourhood O of $\bar{\boldsymbol{y}}$ where $\boldsymbol{H}$ is C^1. Otherwise, let I_z^f and I_z^c be some index sets forming a decomposition of I_z, I_ci^s and I_ci^l sets forming a decomposition of $I_\mathrm{c} \cap I_\mathrm{i}$, and $I_\mathrm{zi}^\mathrm{cs}$, $I_\mathrm{zi}^\mathrm{cl+}$ and $I_\mathrm{zi}^\mathrm{cl-}$ sets forming a further decomposition of $I_\mathrm{z}^\mathrm{c} \cap I_\mathrm{i}$. We associate these decompositions with a function $\boldsymbol{H}^{(i)}: \mathbb{R}^{1+2(n_\Omega+n_D+n_c)} \to \mathbb{R}^{2(n_\Omega+n_D+n_c)}$, $i = i(I_\mathrm{z}^\mathrm{f}, I_\mathrm{z}^\mathrm{c}, I_\mathrm{ci}^\mathrm{s}, I_\mathrm{ci}^\mathrm{l}, I_\mathrm{zi}^\mathrm{cs}, I_\mathrm{zi}^\mathrm{cl+}, I_\mathrm{zi}^\mathrm{cl-})$, defined by

$$
\boldsymbol{H}^{(i)}(\boldsymbol{y}) := \begin{pmatrix}
\boldsymbol{A}(\boldsymbol{u}) - \boldsymbol{L}(\gamma,\boldsymbol{u}) - \boldsymbol{B}_D^\top\boldsymbol{\lambda}_D - \boldsymbol{B}_\nu^\top(\boldsymbol{u})\boldsymbol{\lambda}_\nu - \boldsymbol{B}_\tau^\top(\boldsymbol{u})\boldsymbol{\lambda}_\tau\\
\boldsymbol{B}_D\boldsymbol{u} - \boldsymbol{U}_D(\gamma)\\
\begin{pmatrix} \lambda_{\nu,j}, & j \in I_\mathrm{f} \cup I_\mathrm{z}^\mathrm{f}\\ g_j(\boldsymbol{u}), & j \in I_\mathrm{c} \cup I_\mathrm{z}^\mathrm{c}\end{pmatrix}\\
\begin{pmatrix} \lambda_{\tau,j}, \quad j \in I_\mathrm{f} \cup I_\mathrm{z}^\mathrm{f} \cup I_0\\ (\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j, \quad j \in I_\mathrm{s} \cup I_\mathrm{ci}^\mathrm{s} \cup I_\mathrm{zi}^\mathrm{cs}\\ \lambda_{\tau,j} - s_j\mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u})),\\ \quad j \in ((I_\mathrm{c} \cup I_\mathrm{z}^\mathrm{c}) \cap I_\mathrm{l}) \cup I_\mathrm{ci}^\mathrm{l} \cup I_\mathrm{zi}^\mathrm{cl+} \cup I_\mathrm{zi}^\mathrm{cl-}\end{pmatrix}
\end{pmatrix},
$$

where

$$
s_j = \begin{cases} -\operatorname{sgn}(\bar{\lambda}_{\tau,j} - (\boldsymbol{B}_\tau(\bar{\boldsymbol{u}})\bar{\boldsymbol{u}})_j) \\ \qquad\qquad \text{if } j \in ((I_\mathrm{c} \cup I_\mathrm{z}^\mathrm{c}) \cap I_\mathrm{l}) \cup I_\mathrm{ci}^\mathrm{l},\\ \pm 1 \qquad\qquad \text{if } j \in I_\mathrm{zi}^{\mathrm{cl}\pm}. \end{cases} \tag{43}
$$

Observe that if $\boldsymbol{H}^{(i)}(\boldsymbol{y}) = \boldsymbol{H}(\boldsymbol{y}) = \boldsymbol{0}$ in addition to $\boldsymbol{H}(\bar{\boldsymbol{y}}) = \boldsymbol{0}$, I_z^f and I_z^c correspond to the nodes that are in grazing contact for $\bar{\boldsymbol{y}}$, and not in contact and in contact, respectively, for $\boldsymbol{y}$. Similarly, I_ci^s and I_ci^l correspond to the nodes in strong contact with impending slip for $\bar{\boldsymbol{y}}$, and in stick and slip, respectively, for $\boldsymbol{y}$. Finally, $I_\mathrm{zi}^\mathrm{cs}$, $I_\mathrm{zi}^\mathrm{cl+}$ and $I_\mathrm{zi}^\mathrm{cl-}$ correspond to the nodes in grazing contact with impending slip for $\bar{\boldsymbol{y}}$, and in strong contact with stick, positive and negative slip, respectively, for $\boldsymbol{y}$.

Due to the continuity of the functions $\boldsymbol{u} \mapsto g_j(\boldsymbol{u})$ and $\boldsymbol{u} \mapsto \boldsymbol{B}_\tau(\boldsymbol{u})$, one can find a neighbourhood O of $\bar{\boldsymbol{y}}$ such that

$$
\boldsymbol{H} = \boldsymbol{H}^{(i)} \quad \text{in } D^{(i)},
$$

where

$$
D^{(i)} = \bigcap_{j \in I_\mathrm{z} \cup (I_\mathrm{c} \cap I_\mathrm{i})} D_j^{(i)},
$$

$$
D_j^{(i)} := \begin{cases}
\{\boldsymbol{y} \in O;\ \lambda_{\nu,j} - g_j(\boldsymbol{u}) \ge 0\} & \text{if } j \in I_\mathrm{z}^\mathrm{f},\\
\{\boldsymbol{y} \in O;\ \lambda_{\nu,j} - g_j(\boldsymbol{u}) \le 0\} & \text{if } j \in I_\mathrm{z}^\mathrm{c} \setminus I_\mathrm{i},\\
\{\boldsymbol{y} \in O;\ s_j(\lambda_{\tau,j} - (\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j) \le \mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u}))\} & \\
& \text{if } j \in I_\mathrm{ci}^\mathrm{l},\\
\{\boldsymbol{y} \in O;\ s_j(\lambda_{\tau,j} - (\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j) \ge \mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u}))\} & \\
& \text{if } j \in I_\mathrm{ci}^\mathrm{s},\\
\{\boldsymbol{y} \in O;\ \mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u})) \le \lambda_{\tau,j} - (\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j & \\
\qquad \le -\mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u}))\} & \text{if } j \in I_\mathrm{zi}^\mathrm{cs},\\
\{\boldsymbol{y} \in O;\ \lambda_{\nu,j} - g_j(\boldsymbol{u}) \le 0,\ s_j(\lambda_{\tau,j} - (\boldsymbol{B}_\tau(\boldsymbol{u})\boldsymbol{u})_j) & \\
\qquad \le \mathscr{F}_j(\lambda_{\nu,j} - g_j(\boldsymbol{u}))\} & \text{if } j \in I_\mathrm{zi}^\mathrm{cl+} \cup I_\mathrm{zi}^\mathrm{cl-}
\end{cases}
$$

with s_j defined by (43) and $s_j = -\operatorname{sgn}(\bar{\lambda}_{\tau,j} - (\boldsymbol{B}_\tau(\bar{\boldsymbol{u}})\bar{\boldsymbol{u}})_j)$ if $j \in I_\mathrm{ci}^\mathrm{s}$.

The set $\{\boldsymbol{H}^{(i)}\}_{i \in \mathscr{I}(\bar{\boldsymbol{y}})}$ of selections of $\boldsymbol{H}$ at $\bar{\boldsymbol{y}}$ then consists of $\boldsymbol{H}^{(i)}$ corresponding to all combinations of I_z^f and I_z^c

forming a decomposition of I_z, I_{ci}^s and I_{ci}^l forming a decomposition of $I_c \cap I_i$, and I_{zi}^{cs}, I_{zi}^{cl+} and I_{zi}^{cl-} forming a decomposition of $I_z^c \cap I_i$. In view of the interpretation of the index sets mentioned above, there is a one-to-one correspondence between the selections $\boldsymbol{H}^{(i)}$ and divisions of the contact nodes that are in grazing contact or impending slip for $\bar{\boldsymbol{y}}$ according to the respective contact modes for roots $\boldsymbol{y}$ of $\boldsymbol{H}^{(i)}$ (provided that both $\bar{\boldsymbol{y}}$ and $\boldsymbol{y}$ solve (42)).

Introducing $\boldsymbol{G}^{(i)}$ according to the definition of $D^{(i)}$, one can easily verify that Assumption 1(ii) from the abstract frame is fulfilled. It even holds for $\mathscr{I}'$ from (10) that $\mathscr{I}' = \mathscr{I}(\bar{\boldsymbol{y}})$ for any $\bar{\boldsymbol{y}} \in \mathbb{R}^{1+2(n_\Omega+n_D+n_c)}$. Observe that under Assumption 2, Assumption 3 ensures that in this setting, all solution branches emanating from $\bar{\boldsymbol{y}}$ have *strict* contact modes of all nodes, that is, without any grazing contact or impending slip (see Theorem 2).

Set $n_1 := \#(I_z \cap I_l) + \#(I_c \cap I_i)$ and $n_2 := \#(I_z \cap I_i)$. From the general simplest cases described in Subsection 2.2, the following ones occur in Problem (42):

ad I. $\boldsymbol{L = 2}$: $n_1 = 1, n_2 = 0$ (either one node in grazing contact with non-vanishing slip or one node in strong contact with impending slip);
ad IV. $\boldsymbol{L = 4, M_1 = 1, M_i = 2, i = 2,3,4}$: $n_1 = 0$, $n_2 = 1$ (one node in grazing contact with impending slip);
ad V. $\boldsymbol{L = 4, M_i = 2, i = 1,2,3,4}$: $n_1 = 2, n_2 = 0$ (two nodes in grazing contact with non-vanishing slip or two nodes in strong contact with impending slip or one node in grazing contact with non-vanishing slip and one node in strong contact with impending slip).

To add, the considerations here can be simply modified to the continuation problem proposed in [21, Section 3], which is a bit more general than the parametrised static problem presented here. Application of Theorem 2 to that continuation problem then gives a description of solution branches, which completes Theorem 3, op. cit.

Remark 1 The Dirichlet boundary condition is enforced via a Lagrange multiplier in the present model so that it can be simply parametrised. However, if the Dirichlet condition does not depend on the parameter, it can be prescribed without any significant changes directly in the discrete problem. Let us also note that our abstract frame does not require nodal approximation of the contact conditions either. It covers also discrete problems arising from their integral approximation (after using a numerical quadrature if necessary).

5 Model Examples

In Sub-subsection 5.1.1, the preceding theory will be illustrated on a very simple contact problem, which corresponds to a parametrisation of the example from [18, Section 4] and can be treated analytically. Subsequently, our numerical studies of bifurcations in more realistic models will be presented in Sub-subsection 5.1.2 and Subsection 5.2. The computations for the latter models were performed with the finite-element library GetFEM++ [26]. In particular, Algorithm 1 was used with $\delta_{\max} = 10^{-3}$, $\delta_{\min} = 10^{-6}$, $\delta_{\text{inc}} = 2$ and $\delta_{\text{dec}} = 0.1$, and $|\tau - \tau_0|$ in Step 5 of the algorithm was compared to

$$\tau_{\text{ref}} := 0.02 \max\{|\tau(1) - \tau(0)|, 10^{-8}\}$$

with $\tau(\alpha)$ defined by (37), $\alpha \in \{0,1\}$. The magnitude of $|\tau - \tau_0|$ was considered to be large when it was greater than τ_{ref}, and it was considered as small when it was smaller than $0.5\tau_{\text{ref}}$.

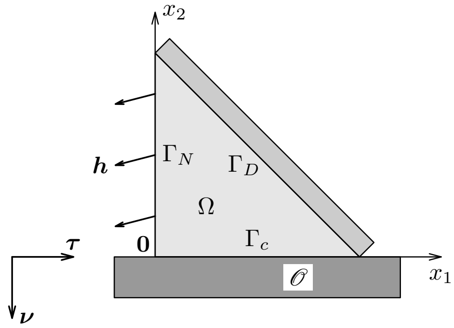


Fig. 11 Geometry of the problem with the triangular body.

5.1 Triangular Body

Let us consider contact of an isosceles triangle with a flat foundation (Fig. 11) in the framework of small-deformation elasticity with Lamé constants λ, $\mu > 0$. The triangle being fixed along Γ_D and volume forces being neglected, the model is parametrised via the surface force $\boldsymbol{h} = \boldsymbol{h}(\gamma) = \gamma(h_1, h_2)$ with h_1 and h_2 constant. The friction coefficient $\mathscr{F} > 0$ is supposed to be constant.

5.1.1 Discretisation with a Single Linear Element

First, we take a model formed by a single linear triangular finite element with the Dirichlet condition on Γ_D prescribed directly so that all degrees of freedom of the model are related to the node in $\mathbf{0}$.

This model leads to Problem (42) with $\boldsymbol{H}\colon \mathbb{R}^5 \to \mathbb{R}^4$, which can be written as

$$\boldsymbol{H}(\boldsymbol{y}) := \begin{pmatrix} au_\nu - bu_\tau + \gamma L_2 - \lambda_\nu \\ -bu_\nu + au_\tau - \gamma L_1 - \lambda_\tau \\ \lambda_\nu - (\lambda_\nu - u_\nu)_- \\ \lambda_\tau - P_{[\mathscr{F}(\lambda_\nu - u_\nu)_-, -\mathscr{F}(\lambda_\nu - u_\nu)_-]}(\lambda_\tau - u_\tau) \end{pmatrix},$$
$$\boldsymbol{y} := (\gamma, u_\nu, u_\tau, \lambda_\nu, \lambda_\tau) \in \mathbb{R}^5,$$

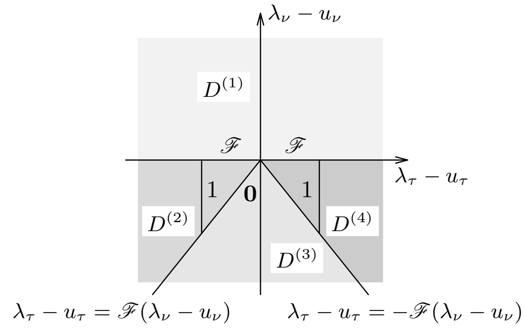


Fig. 12 Structure of the regions for the simple model.

with $a := (\lambda + 3\mu)/2$, $b := (\lambda + \mu)/2$ and the load vector $\boldsymbol{L}(\gamma) = \gamma(L_1, L_2)$.

Let us take $\bar{\boldsymbol{y}} = \boldsymbol{0}$, which corresponds to grazing contact with impending slip of the only contact node, and results thus in the most complex situation. According to the previous section, $\boldsymbol{H} = \boldsymbol{H}^{(i)}$ in $D^{(i)}$, $i \in \mathscr{I}(\bar{\boldsymbol{y}}) = \{1,2,3,4\}$, with

$$D^{(1)} := \{\boldsymbol{y} \in \mathbb{R}^5;\ \lambda_\nu - u_\nu \geq 0\},$$
$$D^{(2)} := \{\boldsymbol{y} \in \mathbb{R}^5;\ \lambda_\nu - u_\nu \leq 0, \lambda_\tau - u_\tau \leq \mathscr{F}(\lambda_\nu - u_\nu)\},$$
$$D^{(3)} := \{\boldsymbol{y} \in \mathbb{R}^5;\ \mathscr{F}(\lambda_\nu - u_\nu) \leq \lambda_\tau - u_\tau \leq -\mathscr{F}(\lambda_\nu - u_\nu)\},$$
$$D^{(4)} := \{\boldsymbol{y} \in \mathbb{R}^5;\ \lambda_\nu - u_\nu \leq 0, -\mathscr{F}(\lambda_\nu - u_\nu) \leq \lambda_\tau - u_\tau\},$$

$$\boldsymbol{H}^{(1)}(\boldsymbol{y}) := \begin{pmatrix} au_\nu - bu_\tau + \gamma L_2 - \lambda_\nu \\ -bu_\nu + au_\tau - \gamma L_1 - \lambda_\tau \\ \lambda_\nu \\ \lambda_\tau \end{pmatrix},$$
$$\boldsymbol{H}^{(2)}(\boldsymbol{y}) := \begin{pmatrix} au_\nu - bu_\tau + \gamma L_2 - \lambda_\nu \\ -bu_\nu + au_\tau - \gamma L_1 - \lambda_\tau \\ u_\nu \\ \mathscr{F}u_\nu - \mathscr{F}\lambda_\nu + \lambda_\tau \end{pmatrix},$$
$$\boldsymbol{H}^{(3)}(\boldsymbol{y}) := \begin{pmatrix} au_\nu - bu_\tau + \gamma L_2 - \lambda_\nu \\ -bu_\nu + au_\tau - \gamma L_1 - \lambda_\tau \\ u_\nu \\ u_\tau \end{pmatrix},$$
$$\boldsymbol{H}^{(4)}(\boldsymbol{y}) := \begin{pmatrix} au_\nu - bu_\tau + \gamma L_2 - \lambda_\nu \\ -bu_\nu + au_\tau - \gamma L_1 - \lambda_\tau \\ u_\nu \\ -\mathscr{F}u_\nu + \mathscr{F}\lambda_\nu + \lambda_\tau \end{pmatrix}$$

(Fig. 12). In this case, Problem ($\mathscr{P}$) coincides with (NF) with $\boldsymbol{A}^{(i)} = \boldsymbol{\nabla H}^{(i)}(\bar{\boldsymbol{y}})$ and $C^{(i)} = D^{(i)}$, $i \in \mathscr{I}' = \{1,2,3,4\}$.

Regarding Assumption 2, one can verify without any difficulties that the gradients $\boldsymbol{\nabla H}^{(i)}(\bar{\boldsymbol{y}})$ have always the full rank for $i = 1,2,3$ whereas $\boldsymbol{\nabla H}^{(4)}(\bar{\boldsymbol{y}})$ is so provided that

$$bL_1 - aL_2 \neq 0 \quad \text{or} \quad \mathscr{F} \neq \frac{a}{b}.$$

Further, Assumption 3 holds if

$$bL_1 - aL_2 \neq 0 \quad \text{and} \quad L_1 \pm \mathscr{F}L_2 \neq 0. \tag{44}$$

This shows that only particular cases are excluded from our general analysis.

Clearly, the branching scenarios in $\bar{\boldsymbol{y}}$ correspond to Case IV from Subsection 2.2. There are always at least two solution rays under satisfaction of (44), and one obtains by elementary calculations that there are even four solution rays if

$$\mathscr{F} > \frac{a}{b} \quad \text{and} \quad (L_1 - \mathscr{F}L_2)(aL_2 - bL_1) > 0.$$

Considering the case

$$\mathscr{F} > \frac{a}{b} \quad \text{and} \quad L_1 > \mathscr{F}L_2 \quad \text{and} \quad aL_2 > bL_1$$

for definiteness (the case $L_1 < \mathscr{F}L_2$ and $aL_2 < bL_1$ being symmetric), one can compute that the solution rays of ($\mathscr{P}$) are generated by

$$\begin{aligned}
\boldsymbol{y}^{(1)} &:= \left(1, \frac{bL_1 - aL_2}{a^2 - b^2}, \frac{aL_1 - bL_2}{a^2 - b^2}, 0, 0\right) && \text{(no contact)}, \\
\boldsymbol{y}^{(2)} &:= \left(-1, 0, -\frac{L_1 + \mathscr{F}L_2}{a + b\mathscr{F}}, \frac{bL_1 - aL_2}{a + b\mathscr{F}}, \mathscr{F}\frac{bL_1 - aL_2}{a + b\mathscr{F}}\right) && \text{(contact-positive slip)}, \\
\boldsymbol{y}^{(3)} &:= (1, 0, 0, L_2, -L_1) && \text{(contact-stick)}, \\
\boldsymbol{y}^{(4)} &:= \left(1, 0, \frac{L_1 - \mathscr{F}L_2}{a - b\mathscr{F}}, \frac{aL_2 - bL_1}{a - b\mathscr{F}}, \mathscr{F}\frac{bL_1 - aL_2}{a - b\mathscr{F}}\right) && \text{(contact-negative slip)},
\end{aligned} \tag{45}$$

$\boldsymbol{y}^{(i)} \in \mathring{D}^{(i)}$. Hence, there is one solution ray with γ negative and there are three solution rays with γ positive. Furthermore,

$$\det\begin{pmatrix} \boldsymbol{\nabla H}^{(1)}(\bar{\boldsymbol{y}}) \\ \boldsymbol{y}^{(1)\top} \end{pmatrix}, \det\begin{pmatrix} \boldsymbol{\nabla H}^{(3)}(\bar{\boldsymbol{y}}) \\ \boldsymbol{y}^{(3)\top} \end{pmatrix} > 0,$$
$$\det\begin{pmatrix} \boldsymbol{\nabla H}^{(2)}(\bar{\boldsymbol{y}}) \\ \boldsymbol{y}^{(2)\top} \end{pmatrix}, \det\begin{pmatrix} \boldsymbol{\nabla H}^{(4)}(\bar{\boldsymbol{y}}) \\ \boldsymbol{y}^{(4)\top} \end{pmatrix} < 0,$$

which implies that the solution ray generated by $\boldsymbol{y}^{(1)}$ is coherently oriented with the ones generated by $\boldsymbol{y}^{(2)}$ and $\boldsymbol{y}^{(4)}$ but incoherently oriented with the one generated by $\boldsymbol{y}^{(3)}$ and so forth.

5.1.2 Discretisation with a Refined Mesh

Next, we consider a problem obtained for a refined mesh of the triangular body to show that the bifurcation behaviour from the simple example preserves in a great extent for problems coming from more realistic discretisations. Namely, the legs of the triangle are considered to be 1 m long and a uniform mesh with 4096 linear triangles and 64 contact nodes is used for the discretisation of the triangle. Further, it is set $\lambda = 100\,\text{GN/m}^2$, $\mu = 82\,\text{GN/m}^2$, $\boldsymbol{h}(\gamma) = \gamma(-26\,\text{GN/m}^2, -7.5\,\text{GN/m}^2)$ and $\mathscr{F} = 1.7$.

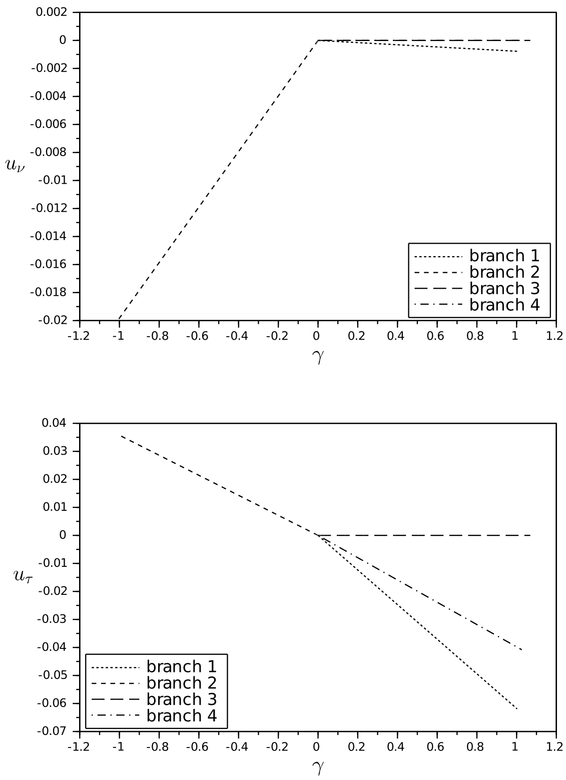


Fig. 14 Bifurcation diagrams.

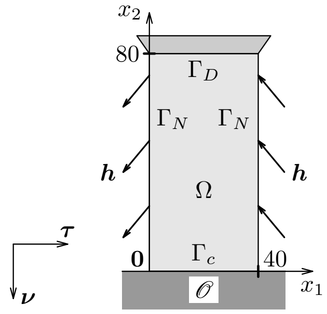


Fig. 16 Geometry of the problem with the rectangular body.

With the aid of the method of piecewise-smooth numerical continuation proposed in [21], we have found four solution branches of the problem corresponding to (42) and emanating from $\bar{\boldsymbol{y}} = \boldsymbol{0}$: one with γ negative and three with γ positive. They correspond to a partial contact and slip of the body to the right, and to no contact, contact-stick and contact-slip to the left of the most left contact node (see Fig. 13). According to the obvious analogy of these branches with the ones from the simple model generated by $\boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(4)}$ from (45), they are denoted as branch 2, branch 1, branch 3 and branch 4, respectively, here and in what follows.

The bifurcation diagrams in Fig. 14 were obtained by plotting the normal and tangential displacements of the most left contact node of the body. One can see from the upper diagram that some components of different branches may coincide (branches 3 and 4 in this case) although the branches do not coincide completely. Especially, this happens for solution components corresponding to the normal displacement (or the x_2-coordinate of the displacement) and to the normal contact stress at the nodes that are in contact and not in contact with the foundation, respectively, for two different branches simultaneously.

When $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ are chosen from various branches, computations with Algorithm 1 show that the numbers of the roots of $\boldsymbol{J}(\alpha)$ defined by (35) in $[0,1]$ vary from zero to two: There is no root for $\boldsymbol{y}_{k-1}$, $\boldsymbol{y}_k$ from the pairs {branch 1, branch 2} and {branch 2, branch 3}, one root for {branch 1, branch 3} and {branch 2, branch 4}, and two roots for {branch 1, branch 4} and {branch 3, branch 4}. Hence, one can conclude that branch 1 is coherently oriented with branches 2 and 4 and incoherently oriented with branch 3 and so forth. Let us emphasise that the orientations remain the same from the simple example although the overall situation is much more complex now – there are 4^{64} regions intersecting at $\bar{\boldsymbol{y}}$ in the present problem!

In the vast majority of our tests, there was no singularity of $\tau(\alpha)$ introduced by (37) in $[0,1]$, that is, no root of $\boldsymbol{M}(\alpha)$ defined by (36) when there was no root of $\boldsymbol{J}(\alpha)$ (see Fig. 15(a) for a typical behaviour of $\tau(\alpha)$ in this case). Further, one and two roots of $\boldsymbol{J}(\alpha)$ were usually closely accompanied by one and two roots of $\boldsymbol{M}(\alpha)$, respectively (Figs. 15(b) and 15(c)).

5.2 Rectangular Body

Finally inspired by [12], we consider contact of a rectangular block that is 40 mm wide and 80 mm high with a flat foundation, see Fig. 16. A plane-strain approximation of the nonlinear Ciarlet-Geymonat constitutive law [8, Chapter 4] is used:

$$\hat{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{F}) = (\tilde{\boldsymbol{\sigma}}(\tilde{\boldsymbol{F}}))_{1 \le i,j \le 2}, \quad \tilde{\boldsymbol{F}} = \begin{pmatrix} \boldsymbol{F} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix}, \quad \boldsymbol{F} \in \mathbb{R}^{2\times 2},$$

$$\tilde{\boldsymbol{\sigma}}(\tilde{\boldsymbol{F}}) = 2b\big(\operatorname{tr}(\tilde{\boldsymbol{F}}^\top \tilde{\boldsymbol{F}})\big)\boldsymbol{I} + 2(a - b\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^\top)\tilde{\boldsymbol{F}} + \big(2c\det(\tilde{\boldsymbol{F}}^\top \tilde{\boldsymbol{F}}) - d\big)\tilde{\boldsymbol{F}}^{-\top}, \quad \tilde{\boldsymbol{F}} \in \mathbb{R}^{3\times 3},$$

where

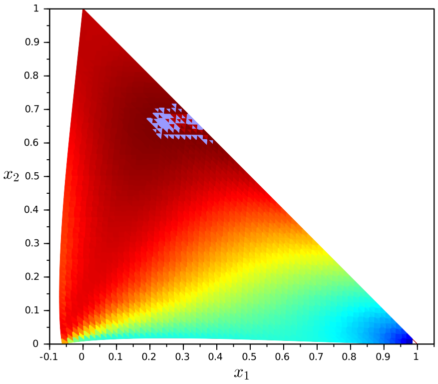
$$\lambda = 4000\ \mathrm{N/mm^2}, \quad \mu = 120\ \mathrm{N/mm^2}, \quad a = 30\ \mathrm{N/mm^2}$$

and

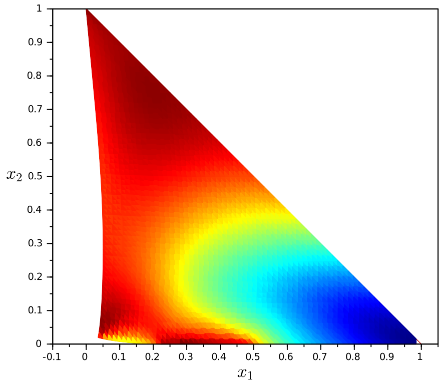
$$b = \frac{\mu}{2} - a, \quad c = \frac{\lambda}{4} - \frac{\mu}{2} + a, \quad d = \frac{\lambda}{2} + \mu.$$

The body is fixed along Γ_D, volume forces are neglected while the surface forces given by the formula

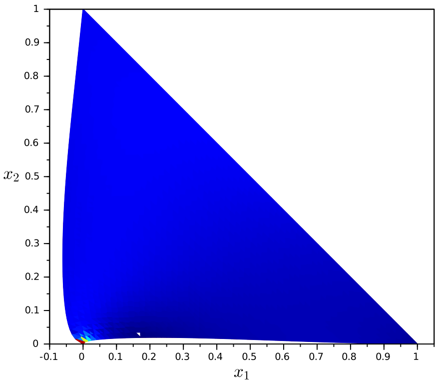
$$\boldsymbol{h}(\boldsymbol{x}, \gamma) = \gamma(-2, 0.12(x_1 - 20))$$



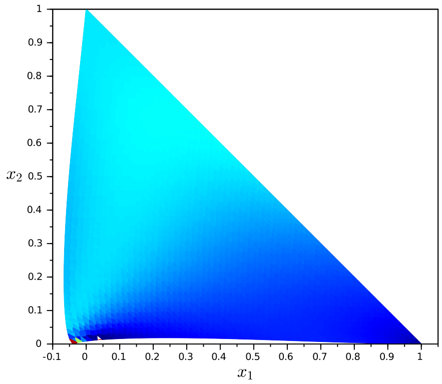
(a) Branch 1.



(b) Branch 2.

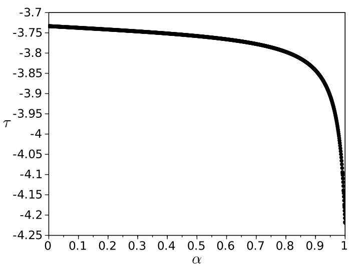


(c) Branch 3.

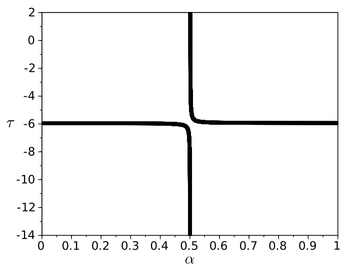


(d) Branch 4.

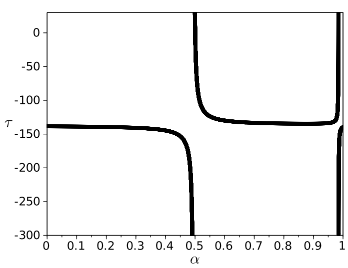
Fig. 13 Solutions with $|\gamma| = 1$: (b) $\gamma = -1$; (a), (c), (d) $\gamma = 1$.



(a) No root of $\boldsymbol{J}(\alpha)$ in $[0,1]$.



(b) One root of $\boldsymbol{J}(\alpha)$ in $[0,1]$.



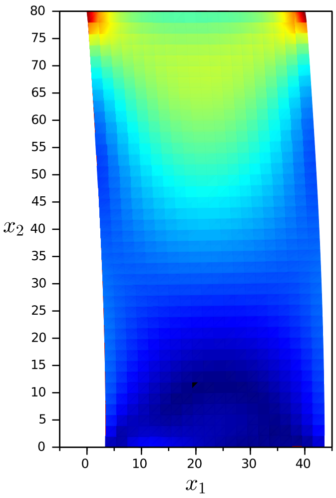
(c) Two roots of $\boldsymbol{J}(\alpha)$ in $[0,1]$.

Fig. 15 Typical behaviours of $\tau(\alpha)$.

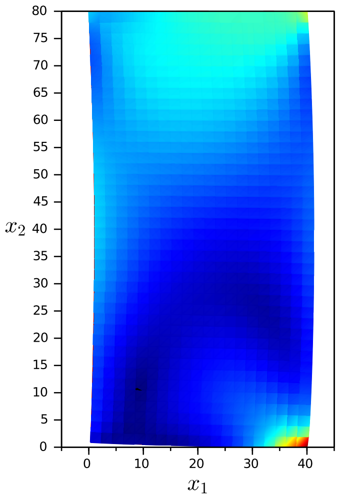
(in $\mathrm{N/mm^2}$) act on both parts of $\varGamma_N$ and $\mathscr{F} = 1$ on $\varGamma_c$. The body is discretised with a uniform mesh with 800 bilinear squares and 21 contact nodes.

We have found six solution branches emanating from $\bar{\boldsymbol{y}} = \boldsymbol{0}$ numerically in this problem: three with γ negative, which correspond to forcing the body to the right and no contact, contact-stick and contact-slip to the right of the most right contact node, and three with γ positive, which correspond to forcing the body to the left and no contact, contact-stick and contact-slip to the left of the most left contact node (Fig. 17).

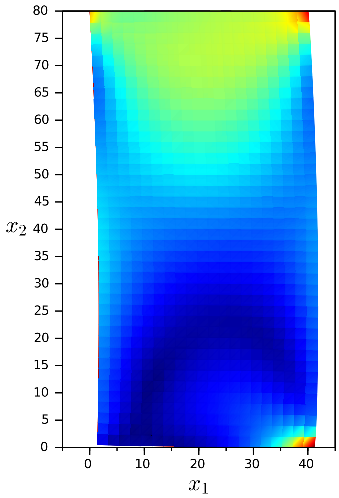
The bifurcation diagrams in Fig. 18 were obtained by plotting the normal and tangential displacements of the most left contact node of the body as previously. In this case, branches 5 and 6 coincide in the upper diagram. One can guess from the lower diagram that some components of some



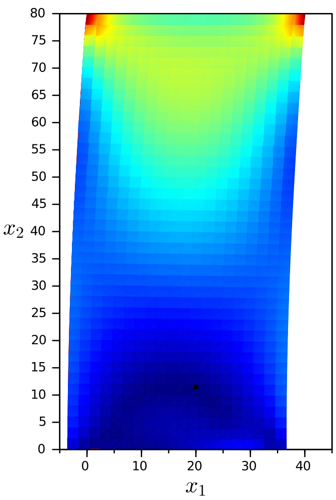
(a) Branch 1.



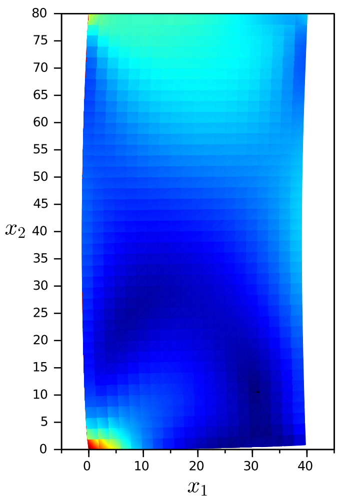
(b) Branch 2.



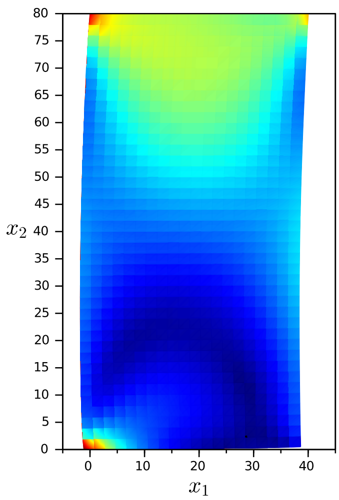
(c) Branch 3.



(d) Branch 4.



(e) Branch 5.



(f) Branch 6.

Fig. 17 Solutions with $|\gamma| = 1$: (a), (b), (c) $\gamma = -1$; (d), (e), (f) $\gamma = 1$.

branches can be continued by the same components of some other branches without loss of differentiability, but the upper diagram clearly shows that this is not the case of all components.

The numbers of the roots of $\boldsymbol{J}(\alpha)$ in $[0,1]$ for various pairs of branches are summarised in Table 1; they vary from zero to two as before. Table 2 presents the resulting orientations of the branches; each branch is coherently oriented with other three branches and incoherently oriented with the other two branches. Behaviours of $\tau(\alpha)$ for various numbers of the roots of $\boldsymbol{J}(\alpha)$ do not differ significantly from Sub-section 5.1.2.

6 Conclusion

The paper presents a complex study of bifurcations in an important class of steady-state piecewise-smooth problems for which the regions of smoothness permit analytical expressions (Assumption 1). Within this class and under certain non-degeneracy assumptions (Assumptions 2 and 3), the study completes the theoretical analysis of local behaviour of the solution set around a non-smooth point from [21] (Theorem 2). It is worth mentioning that apart from being interesting for theoretical study of branching via a simplified problem, our results can also be used for constructing meth-

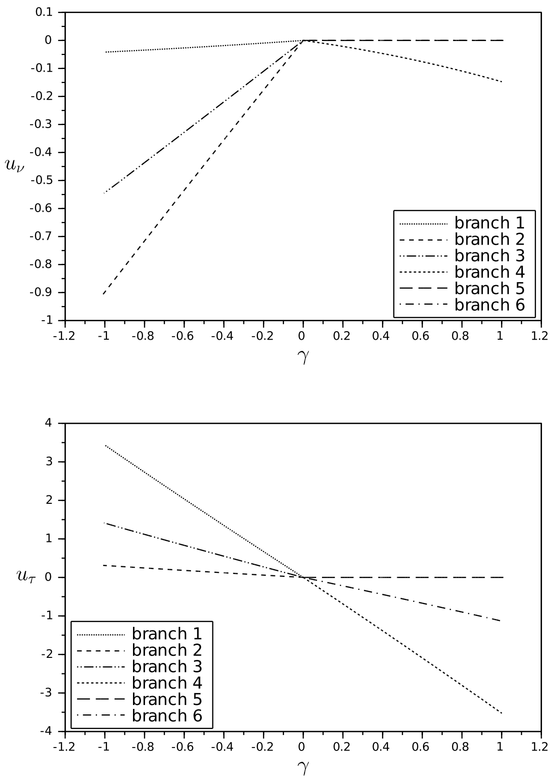


Fig. 18 Bifurcation diagrams.

Table 1 Numbers of the roots of $\boldsymbol{J}(\alpha)$ in $[0,1]$.

branch	1	2	3	4	5	6
1		1	2	0	0	1
2	1		0	0	0	1
3	2	0		1	1	2
4	0	0	1		1	2
5	0	0	1	1		0
6	1	1	2	2	0	

Table 2 Orientations of the branches; the plus and minus signs stand for the coherent and incoherent orientations, respectively.

branch	1	2	3	4	5	6
1		-	+	+	+	-
2	-		+	+	+	-
3	+	+		-	-	+
4	+	+	-		-	+
5	+	+	-	-		+
6	-	-	+	+	+	

ods of numerical continuation of the predictor-corrector type by providing (all possible) tangential predictions.

Furthermore, the most probable branching scenarios have been described and a bifurcation criterion has been formulated. Even though the criterion is based on the particular scenarios, it is applicable generally. Its numerical realisation for large problems has been proposed and tested on plane contact problems with friction. The advantage of the designed algorithm for bifurcation testing is that it does not obey analytical expressions for the regions of smoothness of the PC^1-function involved, and can be thus easily incorporated in a generic continuation routine. Let us emphasise that although the proposed criterion does not detect bifurcations in all possible cases, it is the first attempt to devise such a criterion as far as we know.

All in all, we hope that our contribution illuminates the subject of piecewise-smooth bifurcations and sets up fundamentals for their numerical treatment.

Acknowledgements The basis of this paper was laid down at a post-doctoral stay of T. Ligurský at INSA-Lyon funded by Manufacture Française des Pneumatiques Michelin. During further developments, T. Ligurský received support from the project "Support for building excellent research teams and inter-sectoral mobility at Palacký University in Olomouc II" (CZ.1.07/2.3.00/30.0041).

References

1. Alart, P., Curnier, A.: A mixed formulation for frictional contact problems prone to Newton like solution methods. Comput. Methods Appl. Mech. Engrg. **92**(3), 353–375 (1991)
2. Alexander, J.C.: The topological theory of an embedding method. In: H. Wacker (ed.) Continuation methods, pp. 37–68. Academic Press, New York (1978)
3. Alexander, J.C., Li, T.Y., Yorke, J.A.: Piecewise smooth homotopies. In: B.C. Eaves, F.J. Gould, H.O. Peitgen, M.J. Todd (eds.) Homotopy methods and global convergence, pp. 1–14. Plenum Press, New York (1983)
4. Allgower, E.L., Georg, K.: Numerical path following. In: Handbook of numerical analysis, vol. V, pp. 3–207. North-Holland, Amsterdam (1997)
5. Beremlijski, P., Haslinger, J., Kočvara, M., Outrata, J.V.: Shape optimization in contact problems with Coulomb friction. SIAM J. Optim. **13**(2), 561–587 (2002)
6. di Bernardo, M., Budd, C.J., Champneys, A.R., Kowalczyk, P.: Piecewise-smooth Dynamical Systems: Theory and Applications, *Applied Mathematical Sciences*, vol. 163. Springer-Verlag London (2008)
7. Björkman, G.: Path following and critical points for contact problems. Comput. Mech. **10**(3–4), 231–246 (1992)
8. Ciarlet, P.G.: Mathematical elasticity. Vol. I: Three-dimensional elasticity, *Studies in Mathematics and its Applications*, vol. 20. North-Holland, Amsterdam (1988)
9. Conrad, F., Herbin, R., Mittelmann, H.D.: Approximation of obstacle problems by continuation methods. SIAM J. Numer. Anal. **25**(6), 1409–1431 (1988)
10. Conrad, F., Issard-Roch, F., Brauner, C.M., Nicolaenko, B.: Nonlinear eigenvalue problems in elliptic variational inequalities: a local study. Comm. in Partial Differential Equations **10**(2), 151–190 (1985)
11. Pinto da Costa, A., Martins, J.A.C.: The evolution and rate problems and the computation of all possible evolutions in quasi-static frictional contact. Comput. Methods Appl. Mech. Engrg. **192**(26–27), 2791–2821 (2003)
12. Pinto da Costa, A., Martins, J.A.C.: A numerical study on multiple rate solutions and onset of directional instability in quasi-static frictional contact problems. Computers & Structures **82**(17–19), 1485–1494 (2004)
13. Decker, D.W., Keller, H.B.: Multiple limit point bifurcation. J. Math. Anal. Appl. **75**(2), 417–430 (1980)

14. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol. I. Springer Series in Operations Research. Springer-Verlag, New York (2003)
15. Georg, K.: Matrix-free numerical continuation and bifurcation. Numer. Funct. Anal. Optim. **22**(3–4), 303–320 (2001)
16. Haslinger, J., Janovský, V., Kučera, R.: Path-following the static contact problem with Coulomb friction. In: J. Brandts, S. Korotov, M. Křížek, J. Šístek, T. Vejchodský (eds.) Proceedings of the International Conference Applications of Mathematics 2013, pp. 104–116. Institute of Mathematics, Academy of Sciences of the Czech Republic (2013)
17. Haslinger, J., Janovský, V., Ligurský, T.: Qualitative analysis of solutions to discrete static contact problems with Coulomb friction. Comput. Methods Appl. Mech. Engrg. **205–208**, 149–161 (2012)
18. Hild, P.: On finite element uniqueness studies for Coulomb's frictional contact model. Int. J. Appl. Math. Comput. Sci. **12**(1), 41–50 (2002)
19. Kanno, Y., Martins, J.A.C.: Arc-length method for frictional contact problems using mathematical programming with complementarity constraints. J. Optim. Theory Appl. **131**(1), 89–113 (2006)
20. Klarbring, A.: Stability and critical points in large displacement frictionless contact problems. In: J.A.C. Martins, M. Raous (eds.) Friction and Instabilities, *CISM Courses and Lectures*, vol. 457, pp. 39–64. Springer-Verlag, Wien (2002)
21. Ligurský, T., Renard, Y.: A continuation problem for computing solutions of discretised evolution problems with application to plane quasi-static contact problems with friction. Comput. Methods Appl. Mech. Engrg. **280**, 222–262 (2014)
22. Lundberg, B.N., Poore, A.B.: Numerical continuation and singularity detection methods for parametric nonlinear programming. SIAM J. Optim. **3**(1), 134–154 (1993)
23. Miersemann, E., Mittelmann, H.D.: Continuation for parametrized nonlinear variational inequalities. J. Comput. Appl. Math. **26**(1–2), 23–34 (1989)
24. Paumier, J.C.: Bifurcations et méthodes numériques : applications aux problèmes elliptiques semi-linéaires, *Recherches en Mathématiques Appliquées*, vol. 42. Masson, Paris (1997)
25. Poore, A.B., Tiahrt, C.A.: Bifurcation problems in nonlinear parametric programming. Math. Programming **39**, 189–205 (1987)
26. Renard, Y., Pommier, J.: GetFEM++. An open source generic C++ library for finite element methods, `http://home.gna.org/getfem`
27. Scholtes, S.: Introduction to piecewise differentiable equations. Preprint No. 53/1994, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe (1994)
28. Schulz, M., Pellegrino, S.: Equilibrium paths of mechanical systems with unilateral constraints I. Theory. Proc. R. Soc. Lond. A **456**(2001), 2223–2242 (2000)
29. Tiahrt, C.A., Poore, A.B.: A bifurcation analysis of the nonlinear parametric programming problem. Math. Programming **47**, 117–141 (1990)