



**HAL**  
open science

**Bifurcations in piecewise-smooth steady-state problems:  
abstract study and application to plane contact  
problems with friction**

Tomas Ligursky, Yves Renard

► **To cite this version:**

Tomas Ligursky, Yves Renard. Bifurcations in piecewise-smooth steady-state problems: abstract study and application to plane contact problems with friction. *Computational Mechanics*, 2015, 56 (1), pp.39-62. 10.1007/s00466-015-1154-1 . hal-01077161v2

**HAL Id: hal-01077161**

**<https://hal.science/hal-01077161v2>**

Submitted on 10 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 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 / Revised: March 10, 2015

**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 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 is designed for large problems. 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:

$$\left. \begin{array}{l} \text{Find } \mathbf{y} \in U \text{ such that} \\ \mathbf{H}(\mathbf{y}) = \mathbf{0}, \end{array} \right\} \quad (\mathcal{P})$$

where  $U \subset \mathbb{R}^{N+1}$  and  $\mathbf{H}: U \rightarrow \mathbb{R}^N$ , has been the subject of large number of studies in the last decades (see, e.g., [4, Section 24] and the references therein). If  $\mathbf{H}$  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 the first-order approximation of branches of solutions of variational inequalities were studied in [10, 9, 23], where also methods of numerical continuation of the branches were proposed. 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 were circumvented to a great extent by considering a set of points that are more general than the 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 ( $\mathcal{P}$ ) with  $\mathbf{H}$  piecewise  $C^1$  ( $PC^1$ ) and we developed a method of numerical continuation for this case. We established also some particular results for discretised plane contact problems with Coulomb friction, whose formulation in terms of projections 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, this literature is restricted almost entirely to bifurcations occurring on boundaries between two regions of smooth behaviour of the function involved. However, this setting is not of much interest in steady-state problems as we shall show later on.

The present paper deals with the case of Problem ( $\mathcal{P}$ ) where  $\mathbf{H}$  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 from 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 projection formulation of discretised plane contact problems with Coulomb friction.

The outline of our study is the following: In Section 2, bifurcation points of ( $\mathcal{P}$ ) are analysed in general and a system of piecewise-linear equations capturing the essential features of branching scenarios around certain types of boundary solutions, the so-called border-crossing 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  $\overset{\circ}{A}$  or  $\text{int}A$ , the closure by  $\bar{A}$ ,

the boundary by  $\partial A$ , the exterior by  $\text{ext}A$  and the orthogonal complement by  $A^\perp$ . The gradients of a real-valued function  $f$  and a vector-valued function  $\mathbf{f}$  at a point  $\bar{\mathbf{x}}$  are written as  $\nabla f(\bar{\mathbf{x}})$  and  $\nabla \mathbf{f}(\bar{\mathbf{x}})$ , respectively. If  $\mathbf{f}$  is a function of two variables  $\mathbf{x}$  and  $\mathbf{y}$ ,  $\nabla_{\mathbf{x}} \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  stands for the partial gradient of  $\mathbf{f}$  with respect to  $\mathbf{x}$  at  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ . We use systematically the convention that  $x_i$  and  $f_i$  are the  $i$ th component of a vector  $\mathbf{x}$  and the  $i$ th component function of a vector-valued function  $\mathbf{f}$ , respectively.

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

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

$$\forall \mathbf{y} \in O: \mathbf{H}(\mathbf{y}) \in \{\mathbf{H}^{(i)}(\mathbf{y}); i \in \mathcal{I}(\bar{\mathbf{y}})\}.$$

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

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

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

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

## 2 Theoretical Analysis

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

**Assumption 1** Let  $\mathbf{H}: U \rightarrow \mathbb{R}^N$ ,  $U \subset \mathbb{R}^{N+1}$ , be a  $PC^1$ -function such that every  $\bar{\mathbf{y}} \in U$  meets one of the following two conditions:

- (i)  $\mathbf{H}$  is a  $C^1$ -function in an open neighbourhood  $O \subset U$  of  $\bar{\mathbf{y}}$ ;
- (ii) there exist an open neighbourhood  $O \subset U$  of  $\bar{\mathbf{y}}$  and finite families of  $C^1$ -selections  $\mathbf{H}^{(i)}$  and regions  $D^{(i)}$ ,  $i \in \mathcal{I}(\bar{\mathbf{y}})$ , satisfying

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

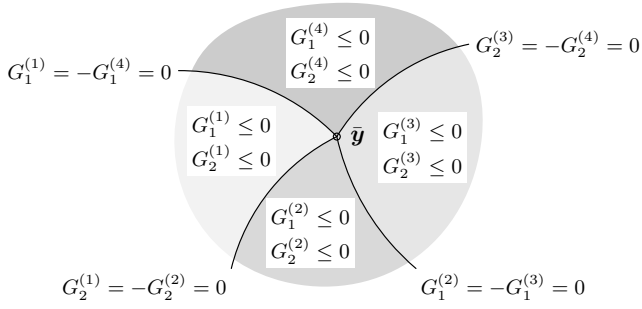
for some  $C^1$ -functions  $\mathbf{G}^{(i)}: O \rightarrow \mathbb{R}^{M_i}$ . Moreover,

$$\mathbf{G}^{(i)}(\bar{\mathbf{y}}) = \mathbf{0}, \quad \forall i \in \mathcal{I}(\bar{\mathbf{y}}), \quad (2)$$

$$\nabla G_j^{(i)}(\bar{\mathbf{y}}) \neq \mathbf{0}, \quad \forall i \in \mathcal{I}(\bar{\mathbf{y}}), j = 1, \dots, M_i, \quad (3)$$

$$\bigcup_{i \in \mathcal{I}(\bar{\mathbf{y}})} D^{(i)} = O, \quad (4)$$

$$\overset{\circ}{D}^{(i)} \cap \overset{\circ}{D}^{(j)} = \emptyset, \quad \forall i, j \in \mathcal{I}(\bar{\mathbf{y}}), i \neq j. \quad (5)$$



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

Let us note that the function class determined by this assumption is very close to the class of functions termed piecewise smooth in [2, Appendix I] and [3]. Expression (1) is natural for  $PC^1$ -functions whose component functions are of the 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  $\mathbf{c}: J \rightarrow \mathbb{R}^{N+1}$  be a curve such that  $\mathbf{H}(\mathbf{c}(s)) = \mathbf{0}$  for any  $s \in J$ . The point  $\mathbf{c}(\bar{s})$  is called a *bifurcation point* of the equation  $\mathbf{H} = \mathbf{0}$  if there exists an  $\varepsilon > 0$  such that every neighbourhood of  $\mathbf{c}(\bar{s})$  contains zero points of  $\mathbf{H}$  that are not on  $\mathbf{c}(\bar{s} - \varepsilon, \bar{s} + \varepsilon)$ .

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

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

The first case leads to well-established theory of smooth bifurcations, whereas the solution set of  $(\mathcal{P})$  is composed of parts of unique solution curves of  $\mathbf{H}^{(i)} = \mathbf{0}$  for individual selections  $\mathbf{H}^{(i)}$  around  $\bar{\mathbf{y}}$  in the second case. The third case can be viewed as a combination of the first two ones; 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  $\bar{\mathbf{y}}$  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  $\nabla \mathbf{H}^{(i)}(\bar{\mathbf{y}})$  of any selection  $\mathbf{H}^{(i)}$ ,  $i \in \mathcal{I}(\bar{\mathbf{y}})$ , has the full rank.

## 2.1 Border-Crossing Normal Form

The analysis of possible branching scenarios is generally carried out by studying a simplified system, which captures the essential features of branching. Following [6, Subsection 3.1.3] and imposing an additional non-degeneracy condition, which will be specified later on, we shall derive such a system for the boundary solution  $\bar{\mathbf{y}}$  in the form of piecewise-linear equations.

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

$$\tilde{\mathbf{H}}(\tilde{\mathbf{y}}) := \mathbf{H}(\tilde{\mathbf{y}} + \bar{\mathbf{y}}), \quad \tilde{\mathbf{H}}^{(i)}(\tilde{\mathbf{y}}) := \mathbf{H}^{(i)}(\tilde{\mathbf{y}} + \bar{\mathbf{y}}),$$

$$\tilde{\mathbf{G}}^{(i)}(\tilde{\mathbf{y}}) := \mathbf{G}^{(i)}(\tilde{\mathbf{y}} + \bar{\mathbf{y}}), \quad \tilde{\mathcal{O}} := \mathcal{O} - \{\bar{\mathbf{y}}\},$$

$$\tilde{D}^{(i)} := \{\tilde{\mathbf{y}} \in \tilde{\mathcal{O}}; \tilde{\mathbf{G}}^{(i)}(\tilde{\mathbf{y}}) \leq \mathbf{0}\}$$

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

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

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

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

$$\tilde{\mathbf{H}}'(\mathbf{0}; \tilde{\mathbf{y}}) = \mathbf{0}. \quad (6)$$

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

$$\mathbf{A}^{(i)} := \nabla \tilde{\mathbf{H}}^{(i)}(\mathbf{0}) = \nabla \mathbf{H}^{(i)}(\bar{\mathbf{y}}), \quad (7)$$

$$\mathbf{B}^{(i)} := \nabla \tilde{\mathbf{G}}^{(i)}(\mathbf{0}) = \nabla \mathbf{G}^{(i)}(\bar{\mathbf{y}}), \quad (8)$$

$$\mathcal{C}^{(i)} := \{\tilde{\mathbf{y}} \in \mathbb{R}^{N+1}; \mathbf{B}^{(i)} \tilde{\mathbf{y}} \leq \mathbf{0}\}, \quad (9)$$

$$\mathcal{I}' := \{i \in \mathcal{I}(\bar{\mathbf{y}}); \mathcal{C}^{(i)} \neq \emptyset\}. \quad (10)$$

**Theorem 1** Under Assumption 1,

$$\bigcup_{i \in \mathcal{I}'} \mathcal{C}^{(i)} = \mathbb{R}^{N+1}, \quad (11)$$

$$\mathcal{C}^{(i)} = \{\tilde{\mathbf{y}} \in \mathbb{R}^{N+1}; \mathbf{B}^{(i)} \tilde{\mathbf{y}} < \mathbf{0}\}, \quad \forall i \in \mathcal{I}', \quad (12)$$

$$\mathcal{C}^{(i)} \cap \mathcal{C}^{(j)} = \emptyset, \quad \forall i, j \in \mathcal{I}', i \neq j, \quad (13)$$

$$\tilde{\mathbf{H}}'(\mathbf{0}; \tilde{\mathbf{y}}) = \mathbf{A}^{(i)} \tilde{\mathbf{y}}, \quad \forall \tilde{\mathbf{y}} \in \mathcal{C}^{(i)}, \forall i \in \mathcal{I}'. \quad (14)$$

*Proof* First, we shall show that

$$\bigcup_{i \in \mathcal{I}(\bar{\mathbf{y}})} \mathcal{C}^{(i)} = \mathbb{R}^{N+1}. \quad (15)$$

Let  $\tilde{\mathbf{y}} \in \mathbb{R}^{N+1}$  be arbitrarily chosen. Since  $\tilde{\mathcal{O}}$  is a neighbourhood of  $\mathbf{0}$ , which is covered by  $\{\tilde{D}^{(i)}\}_{i \in \mathcal{I}(\bar{\mathbf{y}})}$  by (4), and

$\mathcal{S}(\bar{\mathbf{y}})$  is finite, there exist  $i_0 \in \mathcal{S}(\bar{\mathbf{y}})$  and  $\{r_n\} \subset \mathbb{R}$ ,  $r_n \rightarrow 0+$ , such that  $r_n \bar{\mathbf{y}} \in \tilde{D}^{(i_0)}$  for any  $n$ , that is,  $\tilde{\mathbf{G}}^{(i_0)}(r_n \bar{\mathbf{y}}) \leq \mathbf{0}$ . From here and (2),

$$\mathbf{B}^{(i_0)} \bar{\mathbf{y}} = \nabla \tilde{\mathbf{G}}^{(i_0)}(\mathbf{0}) \bar{\mathbf{y}} = \lim_{n \rightarrow \infty} \frac{\tilde{\mathbf{G}}^{(i_0)}(r_n \bar{\mathbf{y}}) - \tilde{\mathbf{G}}^{(i_0)}(\mathbf{0})}{r_n} \leq \mathbf{0},$$

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

Now, suppose that  $\bar{\mathbf{y}} \in C^{(i_0)}$  with  $\mathring{C}^{(i_0)} = \emptyset$ . Due to (15),

$$\bar{\mathbf{y}} \in \overline{\bigcup_{i \in \mathcal{S}(\bar{\mathbf{y}}) \setminus \{i_0\}} C^{(i)}} = \bigcup_{i \in \mathcal{S}(\bar{\mathbf{y}}) \setminus \{i_0\}} \overline{C^{(i)}} = \bigcup_{i \in \mathcal{S}(\bar{\mathbf{y}}) \setminus \{i_0\}} C^{(i)},$$

and (11) follows by induction.

Concerning (12), it suffices to verify

$$\mathring{C}^{(i)} \subset \{\bar{\mathbf{y}} \in \mathbb{R}^{N+1}; \mathbf{B}^{(i)} \bar{\mathbf{y}} < \mathbf{0}\}, \quad i \in \mathcal{S}'.$$

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

$$\mathbf{B}_j^{(i)} \bar{\mathbf{z}} \leq 0,$$

$$\mathbf{B}_j^{(i)} (\bar{\mathbf{z}} - \bar{\mathbf{y}}) \leq 0.$$

But this implies

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

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

Next, take  $\bar{\mathbf{y}} \in \mathring{C}^{(i)}$ ,  $i \in \mathcal{S}'$ , and  $r > 0$  sufficiently small. From the mean-value theorem,

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

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

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

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

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

which holds for an arbitrary  $\bar{\mathbf{y}} \in \mathring{C}^{(i)}$ . Since  $\mathring{C}^{(i)} \neq \emptyset$  by the definition of  $\mathcal{S}'$ , and  $C^{(i)}$  is a closed convex, one has  $\mathring{C}^{(i)} =$

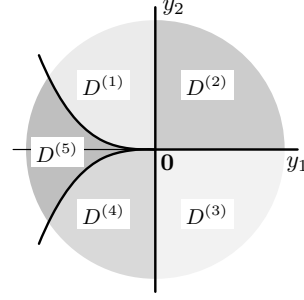


Fig. 2 The regions of smoothness from Example 1.

$C^{(i)}$ . Combining this with the continuity of the function  $\bar{\mathbf{y}} \mapsto \tilde{\mathbf{H}}'(\mathbf{0}; \bar{\mathbf{y}})$ , one arrives at (14).

Finally, let  $\bar{\mathbf{y}} \in \mathring{C}^{(i)} \cap C^{(j)}$ ,  $i, j \in \mathcal{S}'$ ,  $i \neq j$ . In virtue of the equality  $\mathring{C}^{(j)} = C^{(j)}$ , one can find  $\bar{\mathbf{z}} \in \mathring{C}^{(i)} \cap \mathring{C}^{(j)}$ , and arguing as in (16), one gets  $r\bar{\mathbf{z}} \in \tilde{D}^{(i)} \cap \tilde{D}^{(j)}$ . This contradicts to (5).  $\square$

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

*Example 1* Let  $\mathbf{H}: \mathbb{R}^2 \rightarrow \mathbb{R}$  be given by

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}^{(1)}(\mathbf{y}) = -y_1 - y_1^3 - y_2$$

$$\text{in } D^{(1)} = \{\mathbf{y} \in \mathbb{R}^2; -y_1^3 - y_2 \leq 0, y_1 \leq 0\},$$

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}^{(2)}(\mathbf{y}) = -y_2$$

$$\text{in } D^{(2)} = \{\mathbf{y} \in \mathbb{R}^2; -y_1 \leq 0, -y_2 \leq 0\},$$

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}^{(3)}(\mathbf{y}) = y_2$$

$$\text{in } D^{(3)} = \{\mathbf{y} \in \mathbb{R}^2; y_2 \leq 0, -y_1 \leq 0\},$$

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}^{(4)}(\mathbf{y}) = -y_1 - y_1^3 + y_2$$

$$\text{in } D^{(4)} = \{\mathbf{y} \in \mathbb{R}^2; y_1 \leq 0, -y_1^3 + y_2 \leq 0\},$$

$$\mathbf{H}(\mathbf{y}) = \mathbf{H}^{(5)}(\mathbf{y}) = -y_1$$

$$\text{in } D^{(5)} = \{\mathbf{y} \in \mathbb{R}^2; y_1^3 - y_2 \leq 0, y_1^3 + y_2 \leq 0\}$$

(see Fig. 2) and take  $\bar{\mathbf{y}} := \mathbf{0}$ . Then

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

but

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

In the end, we shall show that (6) captures correctly the essential features of the solution set of  $\tilde{\mathbf{H}} = \mathbf{0}$  around  $\mathbf{0}$ , hence of the solution set of  $\mathbf{H} = \mathbf{0}$  around  $\bar{\mathbf{y}}$ . More precisely, we shall see that it determines completely the tangential directions of the solution curves of  $(\mathcal{P})$  emanating from  $\bar{\mathbf{y}}$ . To this end, we shall need the following additional non-degeneracy condition:

**Assumption 3** *Let*

$$\text{Ker } \nabla \mathbf{H}^{(i)}(\bar{\mathbf{y}}) \cap \left( \bigcup_{j=1}^{M_i} \{ \nabla G_j^{(i)}(\bar{\mathbf{y}}) \}^\perp \right) = \{ \mathbf{0} \}, \quad \forall i \in \mathcal{I}(\bar{\mathbf{y}}).$$

Since it is assumed that  $\text{Ker } \nabla \mathbf{H}^{(i)}(\bar{\mathbf{y}})$  are one-dimensional and  $\{ \nabla G_j^{(i)}(\bar{\mathbf{y}}) \}^\perp$  are  $N$ -dimensional subspaces of  $\mathbb{R}^{N+1}$ ,  $i \in \mathcal{I}(\bar{\mathbf{y}})$ ,  $j = 1, \dots, M_i$ , this condition seems to exclude only particular cases. In fact, it 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  $(\mathcal{P})$  from being tangent to the boundaries of the regions  $D^{(i)}$ . As a result, no solution curve of  $(\mathcal{P})$  passing through  $\bar{\mathbf{y}}$  can follow a border between any two regions and thence any approach to any border has to be transversal.

Relaxing a bit the notion from [6], we shall call any boundary solution satisfying Assumption 3 a *border-crossing* solution.

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

$$\mathcal{J}'' := \{ i \in \mathcal{I}(\bar{\mathbf{y}}); \text{Ker } \mathbf{A}^{(i)} \cap \mathring{C}^{(i)} \neq \emptyset \}.$$

*Then the solution set of  $\tilde{\mathbf{H}}'(\mathbf{0}; \bar{\mathbf{y}}) = \mathbf{0}$  coincides either with  $\{ \mathbf{0} \}$  if  $\mathcal{J}'' = \emptyset$ , or with the union  $\bigcup_{i \in \mathcal{J}''} \bigcup_{r \geq 0} r \tilde{\mathbf{y}}^{(i)}$ , where each  $\tilde{\mathbf{y}}^{(i)}$  is arbitrarily chosen from  $\text{Ker } \mathbf{A}^{(i)} \cap \mathring{C}^{(i)}$ .*

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

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

*and the solution set of  $\tilde{\mathbf{H}} = \mathbf{0}$  coincides with the union  $\bigcup_{i \in \mathcal{J}''} \text{Im } \mathbf{c}^{(i)}$  in a vicinity of  $\mathbf{0}$ .*

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

As for the second part, Assumption 1 guarantees that the solution set  $\{ \bar{\mathbf{y}} \in \tilde{O}; \tilde{\mathbf{H}}(\bar{\mathbf{y}}) = \mathbf{0} \}$  is contained in  $\bigcup_{i \in \mathcal{I}(\bar{\mathbf{y}})} \{ \bar{\mathbf{y}} \in \tilde{O}; \tilde{\mathbf{H}}^{(i)}(\bar{\mathbf{y}}) = \mathbf{0} \}$ . So, consider  $i \in \mathcal{I}(\bar{\mathbf{y}})$  fixed.

In virtue of Assumption 2,  $N$  columns of  $\nabla \tilde{\mathbf{H}}^{(i)}(\mathbf{0})$  are linearly independent. For definiteness, we shall consider the case when these are the first  $N$  columns, that is,  $\nabla_{(\tilde{y}_1, \dots, \tilde{y}_N)} \tilde{\mathbf{H}}^{(i)}(\mathbf{0})$  is non-singular, the other cases being analogous. According to the implicit-function theorem,  $\tilde{\mathbf{H}}^{(i)} = \mathbf{0}$  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

$$\mathbf{c}: s \mapsto (\tilde{y}_1(s), \dots, \tilde{y}_N(s), s),$$

one gets from  $\tilde{\mathbf{H}}^{(i)}(\mathbf{c}(s)) = \mathbf{0}$  and the chain rule that

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

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

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

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

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

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

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

Denoting

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

one can find  $\eta_1 > 0$  such that

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

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

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

This together with (18) yields

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

Next, introduce a cone  $\mathcal{C}$  and a real  $\delta' > 0$  by

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

It is readily seen from (21) and Fig. 3 that

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

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

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

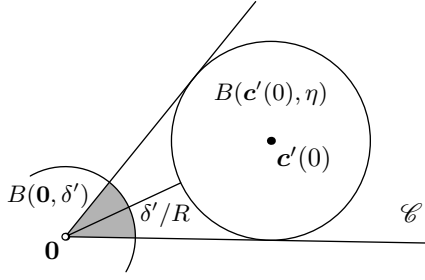


Fig. 3 The intersection  $\mathcal{C} \cap B(\mathbf{0}, \delta')$ .

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

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

and (22) leads to

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

Reducing  $\delta_j$  if necessary, one can show by analogous reasoning that

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

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

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

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

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

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

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

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

$$\mathbf{c}'(0) \in \text{Ker} \mathbf{A}^{(i)} \cap \overset{\circ}{\mathcal{C}}^{(i)},$$

$$\mathbf{c}(s) \in \overset{\circ}{\mathcal{D}}^{(i)} \text{ for any } s \in (0, \delta^{(i)}),$$

$$\mathbf{c}(s) \notin \tilde{\mathcal{D}}^{(i)} \text{ for any } s \in (-\delta^{(i)}, 0).$$

Hence,  $i \in \mathcal{S}''$ ,

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

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

2.  $\nabla^\top \tilde{G}_j^{(i)}(\mathbf{0})\mathbf{c}'(0) > 0$  for any  $j$ . In an analogous way, (17) is satisfied for  $\mathbf{c}^{(i)}: [0, \delta^{(i)}) \ni s \mapsto \mathbf{c}(-s)$  and  $\{\tilde{\mathbf{y}} \in \tilde{\mathcal{O}}; \tilde{\mathbf{H}}(\tilde{\mathbf{y}}) =$

$\mathbf{0}\} \cap \{\tilde{\mathbf{y}} \in \tilde{\mathcal{O}}; \tilde{\mathbf{H}}^{(i)}(\tilde{\mathbf{y}}) = \mathbf{0}\}$  coincides with  $\text{Im} \mathbf{c}^{(i)}$  in a vicinity of  $\mathbf{0}$ .

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

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

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

The proof of the second part of the theorem is finished by getting together the respective cases for all  $i \in \mathcal{S}(\tilde{\mathbf{y}})$ .  $\square$

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

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

where  $\mathbf{F}: \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N$  is a piecewise-linear function defined by

$$\mathbf{F}(\mathbf{y}) := \mathbf{A}^{(i)}\mathbf{y}, \quad \mathbf{y} \in C^{(i)}, \quad i \in \mathcal{S}', \quad (23)$$

with  $\mathbf{A}^{(i)}$ ,  $C^{(i)}$  and  $\mathcal{S}'$  introduced in (7)–(10).

It is worth noticing that (6) is exactly the first-order system  $(\mathcal{P}')$  from [21, Subsection 2.1] shifted to  $\mathbf{0}$ . The relation between  $(\mathcal{P})$  and  $(\mathcal{P}')$  was studied for a general  $PC^1$ -function  $\mathbf{H}$  in [21]. 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  $(\mathcal{P}')$  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  $(\mathcal{P}')$  or of its special case (NF) although these are already simplifications of  $(\mathcal{P})$ . That is why we shall investigate more closely the simplest cases of (NF), which are most likely to occur.

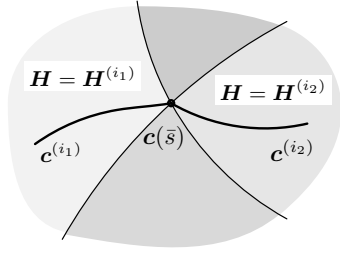
Before that, we shall adapt the notion of orientation from [4]. Let Assumptions 1, 2 and 3 hold and let  $\mathbf{c}^{(i_1)}$  and  $\mathbf{c}^{(i_2)}$  be two  $C^1$ -curves from Theorem 2. Then

$$(\mathbf{c}^{(i)})'_+(0) \in \text{Ker} \nabla \mathbf{H}^{(i)}(\tilde{\mathbf{y}}),$$

or equivalently,  $(\mathbf{c}^{(i)})'_+(0)$  is orthogonal to the  $N$  linearly independent rows of  $\nabla \mathbf{H}^{(i)}(\tilde{\mathbf{y}})$ , and the augmented Jacobian  $\begin{pmatrix} \nabla \mathbf{H}^{(i)}(\tilde{\mathbf{y}}) \\ (\mathbf{c}^{(i)})'_+(0)^\top \end{pmatrix}$  is non-singular for  $i = i_1, i_2$ .

We shall say that the solution curves  $\mathbf{c}^{(i_1)}$  and  $\mathbf{c}^{(i_2)}$  are *coherently oriented* if

$$\det \begin{pmatrix} \nabla \mathbf{H}^{(i_1)}(\tilde{\mathbf{y}}) \\ -(\mathbf{c}^{(i_1)})'_+(0)^\top \end{pmatrix} \det \begin{pmatrix} \nabla \mathbf{H}^{(i_2)}(\tilde{\mathbf{y}}) \\ (\mathbf{c}^{(i_2)})'_+(0)^\top \end{pmatrix} > 0. \quad (24)$$



**Fig. 4** A piecewise-smooth solution curve composed of two smooth curves.

If the converse inequality holds, we shall speak about an *incoherent orientation*.

To justify this concept of orientation, consider the piecewise-smooth curve  $\mathbf{c}: J \rightarrow \mathbb{R}^{N+1}$  defined on an open interval  $J$  containing  $\bar{s}$  by

$$\mathbf{c}(s) = \begin{cases} \mathbf{c}^{(i_1)}(\bar{s} - s) + \bar{\mathbf{y}} & \text{if } s < \bar{s}, \\ \bar{\mathbf{y}} & \text{if } s = \bar{s}, \\ \mathbf{c}^{(i_2)}(s - \bar{s}) + \bar{\mathbf{y}} & \text{if } s > \bar{s} \end{cases} \quad (25)$$

(see Fig. 4). Plainly,  $\mathbf{c}$  is a solution curve of  $(\mathcal{P})$ , and

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

Thence, Condition (24) of coherent orientation can be rewritten in terms of  $\mathbf{c}$  as follows:

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

## 2.2 The Simplest Branching Scenarios

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

$$\begin{aligned} \hat{C}^{(i)} &\neq \emptyset, \quad \text{rank } \mathbf{A}^{(i)} = N, \quad \forall i \in \mathcal{I}', \\ \text{Ker } \mathbf{A}^{(i)} \cap \left( \bigcup_{j=1}^{M_i} \{\mathbf{B}_j^{(i)\top}\}^\perp \right) &= \{\mathbf{0}\}, \quad \forall i \in \mathcal{I}', \end{aligned} \quad (26)$$

where  $\mathbf{B}_j^{(i)}$  stands for the  $j$ th row vector of  $\mathbf{B}^{(i)}$ .

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  $\mathbf{B} \in \mathbb{R}^{(M_1 + \dots + M_L) \times (N+1)}$  as

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

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

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

This will simplify our considerations. Note that after omitting superfluous inequalities if necessary, we may assume that  $\text{rank } \mathbf{B}^{(i)} = M_i$ .

Our study will be based on the following elementary result, which can be immediately used for determination of orientation of two smooth solution curves in the simplest possible case:

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

$$\begin{aligned} \mathbf{F}(\mathbf{y}) &= \mathbf{A}^{(i)}\mathbf{y}, \quad \mathbf{y} \in C^{(i)}, \quad i = 1, 2, \\ C^{(1)} &= \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}^\top \mathbf{y} \leq 0\}, \quad C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}^\top \mathbf{y} \geq 0\} \end{aligned}$$

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

$$\text{rank } \mathbf{A}^{(i)} = N, \quad \text{Ker } \mathbf{A}^{(i)} \cap \{\mathbf{b}\}^\perp = \{\mathbf{0}\}, \quad i = 1, 2. \quad (27)$$

Then for any two vectors  $\mathbf{y}^{(i)} \in \text{Ker } \mathbf{A}^{(i)} \cap \hat{C}^{(i)}$ ,  $i = 1, 2$ ,

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

*Proof* Since  $\mathbf{b} \neq \mathbf{0}$  by (27), there exists an orthonormal basis  $\{\mathbf{q}_1, \dots, \mathbf{q}_{N+1}\}$  of  $\mathbb{R}^{N+1}$  with  $\mathbf{q}_1 = \mathbf{b}/\|\mathbf{b}\|$ , and one can introduce a coordinate transformation  $\mathbf{T}: \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$  by

$$\mathbf{T}: \mathbf{y} \mapsto \hat{\mathbf{y}} := \mathbf{Q}\mathbf{y},$$

where the rows of the matrix  $\mathbf{Q} \in \mathbb{R}^{(N+1) \times (N+1)}$  are formed by  $\mathbf{q}_1^\top, \dots, \mathbf{q}_{N+1}^\top$ . Then one can define a piecewise-linear function  $\hat{\mathbf{F}}: \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N$  by

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

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

$$\hat{C}^{(1)} = \{\hat{\mathbf{y}} \in \mathbb{R}^{N+1}; \hat{y}_1 \leq 0\}, \quad \hat{C}^{(2)} = \{\hat{\mathbf{y}} \in \mathbb{R}^{N+1}; \hat{y}_1 \geq 0\},$$

and one can write

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

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



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

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

it suffices to show that the matrix  $\begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{q}_1^\top \end{pmatrix}$  is non-singular.

But from the imposed assumptions,  $\text{Ker} \mathbf{A}^{(1)}$  is spanned by  $\mathbf{y}^{(1)}$  and  $\mathbf{y}^{(1)} \in \hat{C}^{(1)}$ , which entails that  $\mathbf{q}_1^\top \mathbf{y}^{(1)} < 0$ . Thus,  $\mathbf{q}_1 \notin (\text{Ker} \mathbf{A}^{(1)})^\perp = \text{Im} \mathbf{A}^{(1)\top}$ . Taking into account (27), one deduces that  $\text{rank}(\mathbf{A}^{(1)\top} \mathbf{q}_1) = N + 1$ , that is,  $(\mathbf{A}^{(1)\top} \mathbf{q}_1)$  is non-singular, and so are its transpose as well as  $\hat{\mathbf{A}}_2$ .

Now, taking  $(\hat{y}_1^{(i)}, \hat{y}_2^{(i)}) := \mathbf{T}(\mathbf{y}^{(i)}) = \mathbf{Q}\mathbf{y}^{(i)}$ , one has

$$\begin{aligned} \hat{\mathbf{A}}_1^{(i)} \hat{y}_1^{(i)} + \hat{\mathbf{A}}_2 \hat{y}_2^{(i)} &= \mathbf{0}, \\ \hat{y}_2^{(i)} &= -\hat{y}_1^{(i)} \hat{\mathbf{A}}_2^{-1} \hat{\mathbf{A}}_1^{(i)}. \end{aligned}$$

Here,  $\hat{y}_1^{(1)} < 0$  and  $\hat{y}_1^{(2)} > 0$  as  $\hat{y}^{(1)}$  and  $\hat{y}^{(2)}$  are from  $\hat{C}^{(1)}$  and  $\hat{C}^{(2)}$ , respectively. Finally,

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

because

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

Let us now describe all possible cases of (NF) under the restrictions imposed in the beginning of this subsection.

### 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

$$C^{(1)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}^\top \mathbf{y} \leq 0\}, \quad C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}^\top \mathbf{y} \geq 0\}$$

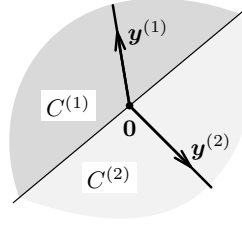


Fig. 5 Solution set in Case I.

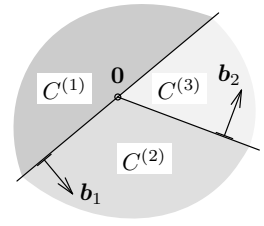


Fig. 6 Structure of the regions in Case II.

for some  $\mathbf{0} \neq \mathbf{b} \in \mathbb{R}^{N+1}$ . As  $\mathbf{b}^\top \mathbf{y} \neq 0$  for any  $\mathbf{y} \in \text{Ker} \mathbf{A}^{(i)} \setminus \{\mathbf{0}\}$  by (26), one can find  $\mathbf{y}^{(i)} \in \text{Ker} \mathbf{A}^{(i)} \cap \hat{C}^{(i)}$ ,  $i = 1, 2$ , and the solution set of (NF) consists of two rays generated by  $\mathbf{y}^{(1)}$  and  $\mathbf{y}^{(2)}$ :

$$\bigcup_{i=1}^2 \bigcup_{r \geq 0} r \mathbf{y}^{(i)}$$

(Fig. 5). Furthermore, direct application of Lemma 1 gives:

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

Regarding our definition of orientation, we shall say that the solution rays are *coherently oriented* in this case.

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

$$F(\mathbf{y}) = (y_1)_+ + y_2 = \begin{cases} y_2 & \text{if } y_1 \leq 0, \\ y_1 + y_2 & \text{if } y_1 \geq 0. \end{cases}$$

### II. $L = 3, M_1 = 1$ .

One can see that the only possibility is:  $M_2 = M_3 = 2$ , and

$$\begin{aligned} C^{(1)} &= \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \leq 0\}, \\ C^{(2)} &= \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_2^\top \mathbf{y} \leq 0\}, \\ C^{(3)} &= \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_2^\top \mathbf{y} \geq 0\} \end{aligned}$$

for some linearly independent vectors  $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^{N+1}$  (see Fig. 6). Due to the continuity of  $\mathbf{F}$ ,

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

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

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

By analogous argumentation, one gets

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

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

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

because  $\mathbf{b}_1$  and  $\mathbf{b}_2$  are linearly independent. In other words, there are only two distinct selections, 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

$$C^{(1)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \leq 0, \mathbf{b}_2^\top \mathbf{y} \leq 0\},$$

$$C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} \leq 0\},$$

$$C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} \geq 0\},$$

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

$$C^{(1)} \cap C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \leq 0, \mathbf{b}_2^\top \mathbf{y} = 0\}, \quad (28)$$

$$C^{(2)} \cap C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} = 0\},$$

$$C^{(1)} \cap C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} = 0, \mathbf{b}_2^\top \mathbf{y} \leq 0\}$$

by the convexity of each  $C^{(i)}$  (Fig. 7(a)).

First, we shall show that (NF) cannot have three distinct solution rays: Taking into account (26), suppose for contradiction that

$$\mathbf{y}^{(i)} \in \text{Ker} \mathbf{A}^{(i)} \cap \overset{\circ}{C}^{(i)}, \quad i = 1, 2, 3.$$

Owing to the continuity of  $\mathbf{F}$  and (28), one has

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

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

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

Hence, defining

$$C^{(12)} := \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \leq 0\},$$

$$C^{(21)} := \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0\},$$

one obtains

$$\mathbf{F}^{(1)} = \mathbf{F}^{(2)} \text{ in } C^{(12)} \cap C^{(21)},$$

and

$$\mathbf{y}^{(1)} \in \overset{\circ}{C}^{(12)}, \quad \mathbf{y}^{(2)} \in \overset{\circ}{C}^{(21)}.$$

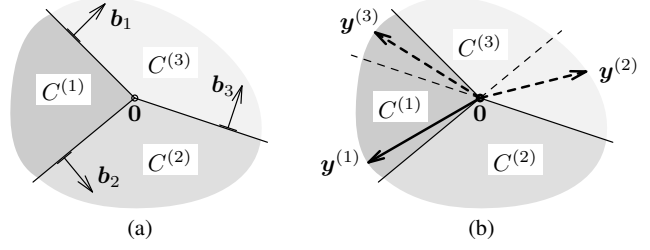
It follows from the application of Lemma 1 to the piecewise-linear function defined to be  $\mathbf{F}^{(1)}$  in  $C^{(12)}$  and  $\mathbf{F}^{(2)}$  in  $C^{(21)}$  that

$$\det \begin{pmatrix} \mathbf{A}^{(1)} \\ -\mathbf{y}^{(1)\top} \end{pmatrix} \det \begin{pmatrix} \mathbf{A}^{(2)} \\ \mathbf{y}^{(2)\top} \end{pmatrix} > 0. \quad (29)$$

By similar reasoning, one gets

$$\det \begin{pmatrix} \mathbf{A}^{(3)} \\ -\mathbf{y}^{(3)\top} \end{pmatrix} \det \begin{pmatrix} \mathbf{A}^{(2)} \\ \mathbf{y}^{(2)\top} \end{pmatrix} > 0, \quad (30)$$

$$\det \begin{pmatrix} \mathbf{A}^{(3)} \\ -\mathbf{y}^{(3)\top} \end{pmatrix} \det \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{y}^{(1)\top} \end{pmatrix} > 0, \quad (31)$$



**Fig. 7** Case III: (a) structure of the regions; (b) a scenario with one solution ray.

and the product of (29), (30) and (31) leads to a contradiction.

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

$$\mathbf{y}^{(1)} \in \text{Ker} \mathbf{A}^{(1)} \cap \overset{\circ}{C}^{(1)}, \quad \text{Ker} \mathbf{A}^{(i)} \cap C^{(i)} = \{\mathbf{0}\}, \quad i = 2, 3.$$

Making use of (26), one can find  $\mathbf{y}^{(2)}$  and  $\mathbf{y}^{(3)}$  such that

$$\mathbf{y}^{(2)} \in \text{Ker} \mathbf{A}^{(2)}, \quad \mathbf{b}_2^\top \mathbf{y}^{(2)} > 0, \quad \mathbf{b}_3^\top \mathbf{y}^{(2)} > 0,$$

$$\mathbf{y}^{(3)} \in \text{Ker} \mathbf{A}^{(3)}, \quad \mathbf{b}_3^\top \mathbf{y}^{(3)} > 0, \quad \mathbf{b}_1^\top \mathbf{y}^{(3)} < 0$$

(so-called *virtual solutions*; see Fig. 7(b)). Taking  $C^{(12)}$  and  $C^{(21)}$  as before and repeating the previous argumentation, one recovers (29). Further, defining

$$C^{(23)} := \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_3^\top \mathbf{y} \leq 0\},$$

$$C^{(32)} := \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_3^\top \mathbf{y} \geq 0\},$$

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

$$\det \begin{pmatrix} \mathbf{A}^{(2)} \\ \mathbf{y}^{(2)\top} \end{pmatrix} \det \begin{pmatrix} \mathbf{A}^{(3)} \\ \mathbf{y}^{(3)\top} \end{pmatrix} > 0. \quad (32)$$

In addition, one gets in a similar way that

$$\det \begin{pmatrix} \mathbf{A}^{(1)} \\ -\mathbf{y}^{(1)\top} \end{pmatrix} \det \begin{pmatrix} \mathbf{A}^{(3)} \\ -\mathbf{y}^{(3)\top} \end{pmatrix} > 0. \quad (33)$$

Comparing (29), (32) and (33), one arrives at a contradiction, again.

On the other hand, examining the scenario with two distinct solution rays, one can conclude that this one is admissible, the two rays being *coherently oriented*. The second admissible scenario consists merely of the trivial solution. Simple examples of these two scenarios follow.

(i) Trivial solution set:

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

(ii) Two solution rays:

$$F(\mathbf{y}) = y_1 + y_2 + P_{[(y_1)_+, +\infty)}(y_2)$$

$$= \begin{cases} y_1 + y_2 & \text{if } y_1 \leq 0, y_2 \leq 0, \\ 2y_1 + y_2 & \text{if } y_1 \geq 0, y_1 \geq y_2, \\ y_1 + 2y_2 & \text{if } y_2 \geq 0, y_2 \geq y_1. \end{cases}$$

**IV.  $L = 4, M_1 = 1$ .**

The only possible situation under our considerations is:  $M_i = 2, i = 2, 3, 4$ , and

$$C^{(1)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \leq 0\},$$

$$C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_2^\top \mathbf{y} \leq 0\},$$

$$C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} \leq 0\},$$

$$C^{(4)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} \geq 0\}$$

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

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

(Fig. 8).

One can show in a similar way as in Case III that the scenarios with exactly one and three solution rays are not possible. On the other hand, the intersection  $\text{Ker} \mathbf{A}^{(1)} \cap \hat{C}^{(1)}$  is always non-empty thanks to (26), that is, (NF) has at least one solution ray. One can conclude that the solution set consists of either two or four solution rays. If there are only two rays, these 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) Two solution rays, which are contained in adjacent cones:

$$F(\mathbf{y}) = 4y_1 + y_2 + 2P_{[-(y_1)_+, (y_1)_+]}(y_2)$$

$$= \begin{cases} 4y_1 + y_2 & \text{if } y_1 \leq 0, \\ 4y_1 + 3y_2 & \text{if } |y_2| \leq y_1, \\ 2y_1 + y_2 & \text{if } y_2 \leq -y_1 \leq 0, \\ 6y_1 + y_2 & \text{if } y_2 \geq y_1 \geq 0. \end{cases}$$

(ii) Two solution rays, which are contained in opposite cones:

$$F(\mathbf{y}) = y_2 + 2P_{[-(y_1)_+, (y_1)_+]}(y_2)$$

$$= \begin{cases} y_2 & \text{if } y_1 \leq 0, \\ 3y_2 & \text{if } |y_2| \leq y_1, \\ y_2 - 2y_1 & \text{if } y_2 \leq -y_1 \leq 0, \\ y_2 + 2y_1 & \text{if } y_2 \geq y_1 \geq 0. \end{cases}$$

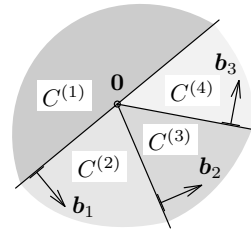


Fig. 8 Case IV.

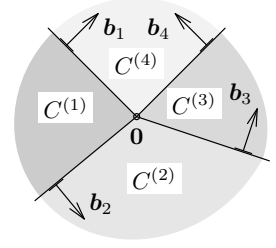


Fig. 9 Case V.

(iii) Four solution rays:

$$F(\mathbf{y}) = -y_2 + 2P_{[-(y_1)_+, (y_1)_+]}(y_2)$$

$$= \begin{cases} -y_2 & \text{if } y_1 \leq 0, \\ y_2 & \text{if } |y_2| \leq y_1, \\ -y_2 - 2y_1 & \text{if } y_2 \leq -y_1 \leq 0, \\ -y_2 + 2y_1 & \text{if } y_2 \geq y_1 \geq 0. \end{cases}$$

**V.  $L = 4, M_i = 2, i = 1, 2, 3, 4$ .**

One can write

$$C^{(1)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \leq 0, \mathbf{b}_2^\top \mathbf{y} \leq 0\},$$

$$C^{(2)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} \leq 0\},$$

$$C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_3^\top \mathbf{y} \geq 0, \mathbf{b}_4^\top \mathbf{y} \leq 0\},$$

$$C^{(4)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} \geq 0, \mathbf{b}_4^\top \mathbf{y} \geq 0\}$$

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

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

$$C^{(2)} \cap C^{(3)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_2^\top \mathbf{y} \geq 0, \mathbf{b}_3^\top \mathbf{y} = 0\},$$

$$C^{(3)} \cap C^{(4)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_3^\top \mathbf{y} \geq 0, \mathbf{b}_4^\top \mathbf{y} = 0\},$$

$$C^{(4)} \cap C^{(1)} = \{\mathbf{y} \in \mathbb{R}^{N+1}; \mathbf{b}_1^\top \mathbf{y} = 0, \mathbf{b}_4^\top \mathbf{y} \geq 0\}.$$

(Fig. 9).

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 consist just of the zero vector.

Examples:

(i) Trivial solution set:

$$F(\mathbf{y}) = -y_1 - y_2 + 2(y_1)_+ + 2(y_2)_+$$

$$= \begin{cases} -y_1 - y_2 & \text{if } y_1 \leq 0, y_2 \leq 0, \\ y_1 - y_2 & \text{if } y_1 \geq 0, y_2 \leq 0, \\ y_1 + y_2 & \text{if } y_1 \geq 0, y_2 \geq 0, \\ -y_1 + y_2 & \text{if } y_1 \leq 0, y_2 \geq 0. \end{cases}$$

(ii) Two solution rays, which are in adjacent cones:

$$F(\mathbf{y}) = y_1 - y_2 + 2(y_1)_+ + 2(y_2)_+$$

$$= \begin{cases} y_1 - y_2 & \text{if } y_1 \leq 0, y_2 \leq 0, \\ 3y_1 - y_2 & \text{if } y_1 \geq 0, y_2 \leq 0, \\ 3y_1 + y_2 & \text{if } y_1 \geq 0, y_2 \geq 0, \\ y_1 + y_2 & \text{if } y_1 \leq 0, y_2 \geq 0. \end{cases}$$

(ii) Two solution rays, which are in opposite cones:

$$F(\mathbf{y}) = y_1 + y_2 + 2(y_1)_+ + 2(y_2)_+$$

$$= \begin{cases} y_1 + y_2 & \text{if } y_1 \leq 0, y_2 \leq 0, \\ 3y_1 + y_2 & \text{if } y_1 \geq 0, y_2 \leq 0, \\ 3y_1 + 3y_2 & \text{if } y_1 \geq 0, y_2 \geq 0, \\ y_1 + 3y_2 & \text{if } y_1 \leq 0, y_2 \geq 0. \end{cases}$$

(ii) Four solution rays:

$$F(\mathbf{y}) = -y_1 + y_2 + 2(y_1)_+ - 2(y_2)_+$$

$$= \begin{cases} -y_1 + y_2 & \text{if } y_1 \leq 0, y_2 \leq 0, \\ y_1 + y_2 & \text{if } y_1 \geq 0, y_2 \leq 0, \\ y_1 - y_2 & \text{if } y_1 \geq 0, y_2 \geq 0, \\ -y_1 - y_2 & \text{if } y_1 \leq 0, y_2 \geq 0. \end{cases}$$

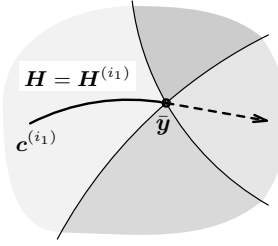
### 2.3 Bifurcation Criterion

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-crossing bifurcation points of Problem ( $\mathcal{P}$ ), which we shall introduce next.

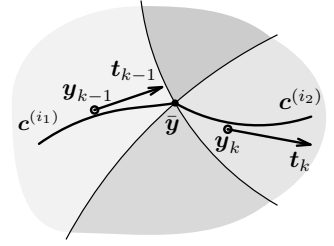
Let Assumptions 1, 2 and 3 hold and let  $\mathbf{c}: J \rightarrow \mathbb{R}^{N+1}$  be a solution curve of ( $\mathcal{P}$ ) defined by (25) for some  $C^1$ -curves  $\mathbf{c}^{(i_1)}$  and  $\mathbf{c}^{(i_2)}$  from Theorem 2. The analysis from the previous subsection suggests us to introduce a **bifurcation criterion** at the border-crossing solution  $\bar{\mathbf{y}}$  of Problem ( $\mathcal{P}$ ) as the condition on *incoherent* orientation of  $\mathbf{c}^{(i_1)}$  and  $\mathbf{c}^{(i_2)}$ . As we have shown in the introduction of our concept of orientation, this condition can be written in view of passing through  $\bar{\mathbf{y}}$  along  $\mathbf{c}$ , namely:

$$\det \begin{pmatrix} \nabla \mathbf{H}^{(i_1)}(\mathbf{c}(\bar{\mathbf{s}})) \\ \mathbf{c}'_{-}(\bar{\mathbf{s}})^{\top} \end{pmatrix} \det \begin{pmatrix} \nabla \mathbf{H}^{(i_2)}(\mathbf{c}(\bar{\mathbf{s}})) \\ \mathbf{c}'_{+}(\bar{\mathbf{s}})^{\top} \end{pmatrix} < 0. \quad (34)$$

In the cases corresponding to the simplest scenarios, this condition is never satisfied when  $\mathbf{c}$  is composed of just two smooth solution branches of ( $\mathcal{P}$ ) emanating from  $\bar{\mathbf{y}}$ . It is fulfilled only if there are four smooth solution branches and  $\mathbf{c}$  is formed by the branches from mutually opposite regions. Besides, consider a smooth solution curve  $\mathbf{c}$  passing through



**Fig. 10** Determination of the region for seeking a new smooth branch.



**Fig. 11** Numerical continuation of  $\mathbf{c}$ .

$\bar{\mathbf{y}}$  and such that  $\mathbf{H}$  is smooth at  $\mathbf{c}(\bar{\mathbf{s}}) = \bar{\mathbf{y}}$ . Then  $\mathbf{c}'_{-}(\bar{\mathbf{s}}) = \mathbf{c}'_{+}(\bar{\mathbf{s}})$ , there is only one selection of  $\mathbf{H}$  in a vicinity of  $\mathbf{c}(\bar{\mathbf{s}})$ , that is,  $\mathbf{H} = \mathbf{H}^{(i_1)} = \mathbf{H}^{(i_2)}$ , and (34) 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  $\mathbf{c}^{(i_1)}$  ending at a boundary solution  $\bar{\mathbf{y}}$  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. 10. 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  $\bar{\mathbf{y}}$  lies in the intersection of four regions. Thus, the potential bifurcation is likely to be detected by the criterion proposed here.

Further, consider that one traces numerically the curve  $\mathbf{c}$  given by (25), namely, that one computes a sequence  $\{\mathbf{y}_k\}$  of points lying approximately on it, and a sequence  $\{\mathbf{t}_k\}$  of approximations of the corresponding tangent vectors. We shall suppose that  $\mathbf{H}$  is smooth at each  $\mathbf{y}_k$ , for simplicity, as it is hardly possible to encounter exactly a point where  $\mathbf{H}$  is not smooth. In addition, let the boundary solution  $\bar{\mathbf{y}}$  lie between  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  such that  $\mathbf{y}_{k-1}, \mathbf{y}_k$  approximate some solutions on  $\mathbf{c}^{(i_1)}$  and  $\mathbf{c}^{(i_2)}$ , respectively (Fig. 11). Assuming that both  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  are close to  $\bar{\mathbf{y}} = \mathbf{c}(\bar{\mathbf{s}})$  and  $\mathbf{t}_{k-1}, \mathbf{t}_k$  are good approximations of  $\mathbf{c}'_{-}(\bar{\mathbf{s}})$  and  $\mathbf{c}'_{+}(\bar{\mathbf{s}})$ , respectively, we arrive from (34) at the following test for detecting border-crossing bifurcations in the course of numerical continuation:

$$\det \begin{pmatrix} \nabla \mathbf{H}(\mathbf{y}_{k-1}) \\ \mathbf{t}_{k-1}^{\top} \end{pmatrix} \det \begin{pmatrix} \nabla \mathbf{H}(\mathbf{y}_k) \\ \mathbf{t}_k^{\top} \end{pmatrix} < 0. \quad (35)$$

Let us note 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 ( $\mathcal{P}$ ) 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 always be straightforward how to determine numerically the sign of the product of the determinants appearing in (35). The aim of this section is to propose a

technique for verification of (35) that requires resolution of a sequence of linear systems instead.

Our approach is based on the following observation, which was employed for smooth bifurcation problems in [15].

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

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

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

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

Then

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

*Proof* The assertion follows directly from Cramer's rule.  $\square$

Next, let  $\mathbf{y}_{k-1}, \mathbf{y}_k, \mathbf{t}_{k-1}$  and  $\mathbf{t}_k$  be given so that both determinants appearing in (35) are non-zero, and let  $\mathbf{b}, \mathbf{c} \in \mathbb{R}^{N+1}$  and  $d \in \mathbb{R}$  be fixed. Set

$$\mathbf{J}(\alpha) := (1 - \alpha) \begin{pmatrix} \nabla \mathbf{H}(\mathbf{y}_{k-1}) \\ \mathbf{t}_{k-1}^\top \end{pmatrix} + \alpha \begin{pmatrix} \nabla \mathbf{H}(\mathbf{y}_k) \\ \mathbf{t}_k^\top \end{pmatrix}, \quad (37)$$

$$\mathbf{M}(\alpha) := \begin{pmatrix} \mathbf{J}(\alpha) & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix} \quad (38)$$

for any  $\alpha \in [0, 1]$ , and whenever  $\mathbf{M}(\alpha)$  is non-singular, take  $\tau(\alpha)$  as a part of the solution of

$$\mathbf{M}(\alpha) \begin{pmatrix} \mathbf{v}(\alpha) \\ \tau(\alpha) \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}. \quad (39)$$

By Lemma 2

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

and observing that

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

one can test (35) equivalently by comparing the signs of  $\det \mathbf{J}(0)$  and  $\det \mathbf{J}(1)$ .

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

$$\det \mathbf{J}(\alpha) = 0 \implies \det \mathbf{M}(\alpha) \neq 0, \quad \alpha \in (0, 1), \quad (40)$$

$$\det \mathbf{M}(0), \det \mathbf{M}(1) \neq 0. \quad (41)$$

Then  $\det \mathbf{M}(\alpha)$  is a non-zero polynomial in  $\alpha$ , and  $\tau(\alpha)$  is a well-defined function with  $\tau(0)$  and  $\tau(1)$  finite. Furthermore, the sign changes of  $\det \mathbf{J}(\alpha)$  are characterised by passings of  $\tau(\alpha)$  through 0 whereas the sign changes of  $\det \mathbf{M}(\alpha)$  by sign changes of  $\tau(\alpha)$  caused by singularities.

Since these two cases can be easily distinguished, we are lead to the following idea for testing (35): To monitor the sign changes of  $\det \mathbf{J}(\alpha)$  by following the behaviour of  $\tau(\alpha)$  when  $\alpha$  passes through  $[0, 1]$ .

It remains to find  $\mathbf{b}, \mathbf{c}$  and  $d$  guaranteeing (40) and (41). Our choice will be based on a classical result [13], [24, Lemma 5.6]:

**Lemma 3** Necessary and sufficient conditions for the non-singularity of  $\mathbf{M} \in \mathbb{R}^{(N+2) \times (N+2)}$  given by (36) are:

- (i)  $d \neq \mathbf{c}^\top \mathbf{J}^{-1} \mathbf{b}$  if  $\mathbf{J}$  is non-singular.
- (ii)  $\dim \text{Ker } \mathbf{J} = 1, \mathbf{b} \notin \text{Im } \mathbf{J}$  and  $\mathbf{c} \notin \text{Im } \mathbf{J}^\top$  if  $\mathbf{J}$  is singular.

As  $\det \mathbf{J}(0)$  and  $\det \mathbf{J}(1)$  are considered to be non-zero,  $\det \mathbf{J}(\alpha)$  is a non-zero polynomial in  $\alpha$  of order at most  $(N+1)$ , and so it has a finite number of roots in  $(0, 1)$  (possibly zero). Denoting them  $\alpha_1, \dots, \alpha_{n_r}$ , one can ensure (40) and (41) by the following assumption:

**Assumption 4** Let

- (i)  $\dim \text{ker } \mathbf{J}(\alpha_i) = 1, \quad i = 1, \dots, n_r;$
- (ii)  $\mathbf{b} \notin \text{Im } \mathbf{J}(\alpha_i), \mathbf{c} \notin \text{Im } \mathbf{J}^\top(\alpha_i), \quad i = 1, \dots, n_r;$
- (iii)  $d \neq \mathbf{c}^\top \mathbf{J}^{-1}(0) \mathbf{b}, \mathbf{c}^\top \mathbf{J}^{-1}(1) \mathbf{b}.$

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

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

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

and setting

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

one has according to the analysis in Subsection 2.2:

$$\text{rank } \mathbf{A}^{(1)} = \text{rank } \mathbf{A}^{(2)} = N,$$

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

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

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

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

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

where the rows of  $\mathbf{Q} \in \mathbb{R}^{(N+1) \times (N+1)}$  are formed by  $\mathbf{q}_1^\top, \dots, \mathbf{q}_{N+1}^\top$ . One can deduce as before that

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

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

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

and consequently

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

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

$$N \leq \text{rank} \begin{pmatrix} (1-\alpha_i)\nabla\mathbf{H}(\mathbf{y}_{k-1}) + \alpha_i\nabla\mathbf{H}(\mathbf{y}_k) \\ (1-\alpha_i)\mathbf{t}_{k-1}^\top + \alpha_i\mathbf{t}_k^\top \end{pmatrix} = \text{rank} \mathbf{J}(\alpha_i),$$

$$i = 1, \dots, n_r.$$

This shows satisfaction of Assumption 4(i) for  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  from adjacent regions.

Secondly, suppose that  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  belong to solution branches from opposite regions that meet at  $\bar{\mathbf{y}}$  and correspond to Case IV or V. Let

$$\begin{aligned} \nabla\mathbf{H}(\mathbf{y}_{k-1}) &= \nabla\mathbf{H}^{(1)}(\mathbf{y}_{k-1}), \quad \nabla\mathbf{H}(\mathbf{y}_k) = \nabla\mathbf{H}^{(3)}(\mathbf{y}_k), \\ \mathbf{A}^{(i)} &= \nabla\mathbf{H}^{(i)}(\bar{\mathbf{y}}), \quad i = 1, 2, 3. \end{aligned}$$

Then

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

and, similarly as before, one can find linearly independent  $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^{N+1}$  such that

$$\begin{aligned} \mathbf{A}^{(1)}\mathbf{y} &= \mathbf{A}^{(2)}\mathbf{y}, \quad \forall \mathbf{y} \in \mathbb{R}^{N+1} \text{ with } \mathbf{b}_1^\top \mathbf{y} = 0, \\ \mathbf{A}^{(2)}\mathbf{y} &= \mathbf{A}^{(3)}\mathbf{y}, \quad \forall \mathbf{y} \in \mathbb{R}^{N+1} \text{ with } \mathbf{b}_2^\top \mathbf{y} = 0, \\ \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{b}_1^\top \end{pmatrix} &\text{ is non-singular.} \end{aligned}$$

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

$$\hat{\mathbf{A}}^{(i)} := \mathbf{A}^{(i)}\mathbf{Q}^{-1}, \quad i = 1, 2, 3. \quad (42)$$

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

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

Consequently, one can show that

$$\begin{aligned} \hat{\mathbf{A}}^{(1)} &= (\hat{\mathbf{A}}_1^{(1)}, \hat{\mathbf{A}}_2^{(1)}, \hat{\mathbf{A}}_3), \quad \hat{\mathbf{A}}^{(2)} = (\hat{\mathbf{A}}_1^{(3)}, \hat{\mathbf{A}}_2^{(1)}, \hat{\mathbf{A}}_3), \\ \hat{\mathbf{A}}^{(3)} &= (\hat{\mathbf{A}}_1^{(3)}, \hat{\mathbf{A}}_2^{(3)}, \hat{\mathbf{A}}_3) \end{aligned}$$

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

$$N-1 \leq \text{rank} \begin{pmatrix} (1-\alpha_i)\nabla\mathbf{H}(\mathbf{y}_{k-1}) + \alpha_i\nabla\mathbf{H}(\mathbf{y}_k) \\ (1-\alpha_i)\mathbf{t}_{k-1}^\top + \alpha_i\mathbf{t}_k^\top \end{pmatrix} = \text{rank} \mathbf{J}(\alpha_i),$$

$$i = 1, \dots, n_r,$$

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

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

$$\begin{aligned} (1-\alpha)\hat{\mathbf{A}}_j^{(1)} + \alpha\hat{\mathbf{A}}_j^{(3)} &\in \text{Im} \hat{\mathbf{A}}_3, \\ \hat{\mathbf{A}}_j^{(3)} &\in \frac{\alpha-1}{\alpha}\hat{\mathbf{A}}_j^{(1)} + \text{Im} \hat{\mathbf{A}}_3 \end{aligned}$$

for  $j = 1, 2$ . By (42),

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

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

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

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

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

Last, if  $\mathbf{b}$  and  $\mathbf{c}$  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  $\mathbf{b}$ ,  $\mathbf{c}$  and  $d$ , so one can suppose Assumption 4 to hold when choosing them randomly. We propose the following numerical procedure for testing Condition (35):

#### Algorithm 1

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

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

Step 2: Set

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

Step 3: Solve

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

for  $\tau$ .

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  $n_{\text{ch}}$  is odd, print ‘‘bifurcation detected’’.

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

The values of  $\alpha$  are increased by the current values of  $\delta$ . The latter one is bounded by  $\delta_{\min}$  and  $\delta_{\max}$  and adapted by the scale factors  $\delta_{\text{inc}}$  and  $\delta_{\text{dec}}$  according to the latest values of  $\tau$ . This adaptivity serves for effective treatment of singularities of  $\tau$ , which are characterised by large  $|\tau - \tau_0|$ . Let us point out that one has to choose  $\delta_{\min}$  and  $\delta_{\text{dec}}$  small enough to detect correctly all singularities in computations. On the other hand, it is highly improbable to encounter exactly a value of  $\alpha$  with  $\mathbf{M}(\alpha)$  singular and so 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  $\Gamma_D$ ,  $\Gamma_N$  and  $\Gamma_c$ . Denoting the deformation of the body by  $\boldsymbol{\varphi}$ , we can express it in terms of the displacement  $\mathbf{u}$  as  $\boldsymbol{\varphi} = \text{id} + \mathbf{u}$ , where  $\text{id}$  stands for the identity mapping.

The displacement  $\mathbf{u}_D$  is imposed on  $\Gamma_D$ , the body in the deformed configuration  $\boldsymbol{\varphi}(\Omega)$  is subject to the body forces of the density  $\mathbf{f}^\boldsymbol{\varphi}$ , and the surface forces of the density  $\mathbf{h}^\boldsymbol{\varphi}$  act on  $\boldsymbol{\varphi}(\Gamma_N)$ . We suppose that the prescribed displacement and the forces may depend on a real parameter  $\gamma$ , in general, that is,  $\mathbf{u}_D = \mathbf{u}_D(\mathbf{x}, \gamma)$ ,  $\mathbf{f}^\boldsymbol{\varphi} = \mathbf{f}^\boldsymbol{\varphi}(\mathbf{x}^\boldsymbol{\varphi}, \gamma)$  and  $\mathbf{h}^\boldsymbol{\varphi} = \mathbf{h}^\boldsymbol{\varphi}(\mathbf{x}^\boldsymbol{\varphi}, \gamma)$ , where  $\mathbf{x}$  and  $\mathbf{x}^\boldsymbol{\varphi}$  stand for points in the initial and the deformed configuration, respectively.

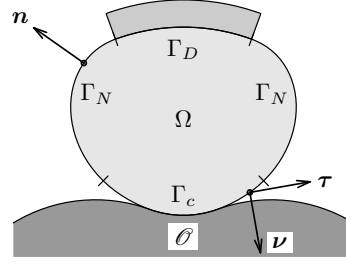


Fig. 12 Geometry of the problem.

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

$$\begin{aligned} g(\mathbf{x}) &> 0, & \forall \mathbf{x} \in V \cap \text{int } \mathcal{O}, \\ g(\mathbf{x}) &= 0, & \forall \mathbf{x} \in \partial\mathcal{O}, \\ g(\mathbf{x}) &< 0, & \forall \mathbf{x} \in V \cap \text{ext } \mathcal{O}. \end{aligned}$$

As a consequence, one can extend the unit inward normal  $\mathbf{v}$  and the unit tangent  $\boldsymbol{\tau}$  to the obstacle from  $\partial\mathcal{O}$  to  $V$  as

$$\mathbf{v}(\mathbf{x}) = \frac{\nabla g(\mathbf{x})}{\|\nabla g(\mathbf{x})\|}, \quad \boldsymbol{\tau}(\mathbf{x}) = (-v_2(\mathbf{x}), v_1(\mathbf{x}))$$

(see Fig. 12).

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

$$\left. \begin{aligned} &\text{Find } \mathbf{u} \in \mathcal{U}_{\text{ad}} \text{ such that} \\ &\mathbf{div } \boldsymbol{\sigma}(\mathbf{x}) + \mathbf{f}(\mathbf{x}, \gamma) = \mathbf{0}, \quad \mathbf{x} \in \Omega, \\ &\boldsymbol{\sigma}(\mathbf{x}) = \hat{\boldsymbol{\sigma}}(\mathbf{x}, \mathbf{I} + \nabla \mathbf{u}(\mathbf{x})), \quad \mathbf{x} \in \Omega, \\ &\mathbf{u}(\mathbf{x}) = \mathbf{u}_D(\mathbf{x}, \gamma), \quad \mathbf{x} \in \Gamma_D, \\ &\boldsymbol{\sigma}(\mathbf{x})\mathbf{n}(\mathbf{x}) = \mathbf{h}(\mathbf{x}, \gamma), \quad \mathbf{x} \in \Gamma_N, \\ &\left. \begin{aligned} g(\mathbf{x} + \mathbf{u}(\mathbf{x})) &\leq 0, \quad T_v(\mathbf{x}) \leq 0, \\ g(\mathbf{x} + \mathbf{u}(\mathbf{x}))T_v(\mathbf{x}) &= 0, \end{aligned} \right\} \mathbf{x} \in \Gamma_c, \\ &\left. \begin{aligned} |T_\tau(\mathbf{x})| &\leq -\mathcal{F}(\mathbf{x})T_v(\mathbf{x}), \\ u_\tau(\mathbf{x}) \neq 0 &\implies T_\tau(\mathbf{x}) = \mathcal{F}(\mathbf{x})T_v(\mathbf{x}) \frac{u_\tau(\mathbf{x})}{|u_\tau(\mathbf{x})|}, \end{aligned} \right\} \mathbf{x} \in \Gamma_c, \end{aligned} \right\} (43)$$

where  $\gamma$  varies over an interval of interest. The set  $\mathcal{U}_{\text{ad}}$  of kinematically admissible displacements is introduced as

$$\mathcal{U}_{\text{ad}} := \{\mathbf{v}: \overline{\Omega} \rightarrow \mathbb{R}^2 \text{ ‘‘smooth enough’’};$$

$$\text{id} + \mathbf{v} \text{ is injective in } \Omega, \det(\mathbf{I} + \nabla \mathbf{v}) > 0 \text{ in } \overline{\Omega}\},$$

$\boldsymbol{\sigma}$  denotes the first Piola-Kirchhoff stress tensor,  $\hat{\boldsymbol{\sigma}}$  its response function characterising the material of the elastic body

and  $\mathbf{I}$  is the identity matrix. Further,  $\mathbf{n}$  stands for the unit outward normal vector along  $\partial\Omega$  and

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

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

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

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

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

is the tangential displacement.

Discretisation of this problem is done by applying a Lagrange finite-element method to a mixed variational formulation of (43) with Lagrange multipliers enforcing the Dirichlet and the contact boundary conditions. In particular, we consider nodal approximation of the contact conditions written in terms of projections as proposed in [1]. This leads to a discrete problem in the form of ( $\mathcal{P}$ ):

$$\left. \begin{aligned} \text{Find } \mathbf{y} &:= (\gamma, \mathbf{u}, \boldsymbol{\lambda}_D, \boldsymbol{\lambda}_v, \boldsymbol{\lambda}_\tau) \in \mathbb{R}^{1+2(n_\Omega+n_D+n_c)} \\ \text{such that} \\ \mathbf{H}(\mathbf{y}) &= \mathbf{0}, \end{aligned} \right\} \quad (44)$$

where  $\mathbf{H}: \mathbb{R}^{1+2(n_\Omega+n_D+n_c)} \rightarrow \mathbb{R}^{2(n_\Omega+n_D+n_c)}$  is defined by

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

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

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

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

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

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

$$\begin{aligned} I_f &:= \{j \in I; \bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}) > 0\}, \\ I_c &:= \{j \in I; \bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}) < 0\}, \\ I_z &:= \{j \in I; \bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}) = 0\}, \\ I_0 &:= \{j \in I; \mathcal{F}_j = 0\}, \\ I_s &:= \{j \in I; \\ &\quad \mathcal{F}_j > 0, |\bar{\lambda}_{\tau,j} - (\mathbf{B}_\tau(\bar{\mathbf{u}})\bar{\mathbf{u}})_j| < -\mathcal{F}_j(\bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}))_-\}, \\ I_l &:= \{j \in I; \\ &\quad \mathcal{F}_j > 0, |\bar{\lambda}_{\tau,j} - (\mathbf{B}_\tau(\bar{\mathbf{u}})\bar{\mathbf{u}})_j| > -\mathcal{F}_j(\bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}))_-\}, \\ I_i &:= \{j \in I; \\ &\quad \mathcal{F}_j > 0, |\bar{\lambda}_{\tau,j} - (\mathbf{B}_\tau(\bar{\mathbf{u}})\bar{\mathbf{u}})_j| = -\mathcal{F}_j(\bar{\lambda}_{v,j} - g_j(\bar{\mathbf{u}}))_-\}. \end{aligned}$$

Apparently, if  $\bar{\mathbf{y}}$  is a solution of (44),  $I_f$  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  $\mathbf{u} \mapsto g_j(\mathbf{u})$  and  $\mathbf{u} \mapsto \mathbf{B}_\tau(\mathbf{u})$  are continuous. It is thus readily seen that if  $I_z \cup (I_c \cap I_l) = \emptyset$ , there exists a neighbourhood  $O$  of  $\bar{\mathbf{y}}$  where  $\mathbf{H}$  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  $I_{ci}^l$  index sets forming a decomposition of  $I_c \cap I_l$ , and  $I_{zi}^{cs}$ ,  $I_{zi}^{cl+}$  and  $I_{zi}^{cl-}$  index sets forming a further decomposition of  $I_z^c \cap I_l$ . We associate these decompositions with a function  $\mathbf{H}^{(i)}: \mathbb{R}^{1+2(n_\Omega+n_D+n_c)} \rightarrow \mathbb{R}^{2(n_\Omega+n_D+n_c)}$ ,  $i = i(I_z^f, I_z^c, I_{ci}^s, I_{ci}^l, I_{zi}^{cs}, I_{zi}^{cl+}, I_{zi}^{cl-})$ , defined by

$$\mathbf{H}^{(i)}(\mathbf{y}) := \begin{pmatrix} \mathbf{A}(\mathbf{u}) - \mathbf{L}(\gamma, \mathbf{u}) - \mathbf{B}_D^\top \boldsymbol{\lambda}_D - \mathbf{B}_v^\top(\mathbf{u}) \boldsymbol{\lambda}_v - \mathbf{B}_\tau^\top(\mathbf{u}) \boldsymbol{\lambda}_\tau \\ \mathbf{B}_D \mathbf{u} - \mathbf{U}_D(\gamma) \\ \left( \begin{array}{l} \lambda_{v,j}, \quad j \in I_f \cup I_z^f \\ g_j(\mathbf{u}), \quad j \in I_c \cup I_z^c \end{array} \right) \\ \left( \begin{array}{l} \lambda_{\tau,j}, \quad j \in I_f \cup I_z^f \cup I_0 \\ (\mathbf{B}_\tau(\mathbf{u})\mathbf{u})_j, \quad j \in I_s \cup I_{ci}^s \cup I_{zi}^{cs} \\ \lambda_{\tau,j} - s_j \mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u})), \\ \quad j \in ((I_c \cup I_z^c) \cap I_l) \cup I_{ci}^l \cup I_{zi}^{cl+} \cup I_{zi}^{cl-} \end{array} \right) \end{pmatrix},$$



where

$$s_j = \begin{cases} -\operatorname{sgn}(\bar{\lambda}_{\tau,j} - (\mathbf{B}_\tau(\bar{\mathbf{u}})\bar{\mathbf{u}})_j) & \text{if } j \in ((I_c \cup I_z^c) \cap I_i) \cup I_{ci}^l, \\ \pm 1 & \text{if } j \in I_{zi}^{\text{cl}\pm}. \end{cases} \quad (45)$$

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

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

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

where

$$D^{(i)} := \bigcap_{j \in I_z \cup (I_c \cap I_i)} D_j^{(i)},$$

$$D_j^{(i)} = \begin{cases} \{\mathbf{y} \in O; \lambda_{v,j} - g_j(\mathbf{u}) \geq 0\} & \text{if } j \in I_z^f, \\ \{\mathbf{y} \in O; \lambda_{v,j} - g_j(\mathbf{u}) \leq 0\} & \text{if } j \in I_z^c \setminus I_i, \\ \{\mathbf{y} \in O; s_j(\lambda_{\tau,j} - (\mathbf{B}_\tau(\mathbf{u})\mathbf{u})_j) \leq \mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u}))\} & \text{if } j \in I_{ci}^l, \\ \{\mathbf{y} \in O; s_j(\lambda_{\tau,j} - (\mathbf{B}_\tau(\mathbf{u})\mathbf{u})_j) \geq \mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u}))\} & \text{if } j \in I_{ci}^s, \\ \{\mathbf{y} \in O; \mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u})) \leq \lambda_{\tau,j} - (\mathbf{B}_\tau(\mathbf{u})\mathbf{u})_j \\ \leq -\mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u}))\} & \text{if } j \in I_{zi}^{\text{cs}}, \\ \{\mathbf{y} \in O; \lambda_{v,j} - g_j(\mathbf{u}) \leq 0, s_j(\lambda_{\tau,j} - (\mathbf{B}_\tau(\mathbf{u})\mathbf{u})_j) \\ \leq \mathcal{F}_j(\lambda_{v,j} - g_j(\mathbf{u}))\} & \text{if } j \in I_{zi}^{\text{cl}+} \cup I_{zi}^{\text{cl}-} \end{cases}$$

with  $s_j$  defined by (45) and  $s_j = -\operatorname{sgn}(\bar{\lambda}_{\tau,j} - (\mathbf{B}_\tau(\bar{\mathbf{u}})\bar{\mathbf{u}})_j)$  if  $j \in I_{ci}^s$ .

The set  $\{\mathbf{H}^{(i)}\}_{i \in \mathcal{I}(\bar{\mathbf{y}})}$  of selections of  $\mathbf{H}$  at  $\bar{\mathbf{y}}$  then consists of  $\mathbf{H}^{(i)}$  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  $I_{zi}^{\text{cs}}$ ,  $I_{zi}^{\text{cl}+}$  and  $I_{zi}^{\text{cl}-}$  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  $\mathbf{H}^{(i)}$  and divisions of the contact nodes that are in grazing contact or impending slip for  $\bar{\mathbf{y}}$  according to the respective contact modes for  $\mathbf{y}$  provided that  $\mathbf{H}(\bar{\mathbf{y}}) = \mathbf{0} = \mathbf{H}(\mathbf{y}) = \mathbf{H}^{(i)}(\mathbf{y})$ .

Introducing  $\mathbf{G}^{(i)}$  according to the definition of  $D^{(i)}$  above, one can easily verify that Assumption 1(ii) from the abstract frame is fulfilled. One can even show that  $\mathcal{S}' = \mathcal{S}(\bar{\mathbf{y}})$  for any  $\bar{\mathbf{y}} \in \mathbb{R}^{1+2(n_\Omega+n_D+n_c)}$  for  $\mathcal{S}'$  given by (10). Moreover, Assumption 3 ensures under Assumption 2 that, in this setting, all solution branches emanating from  $\bar{\mathbf{y}}$  have *strict* contact

modes of all nodes, that is, without any grazing contact or impending slip (compare to [21, Remark 6]).

Next, set  $n_1 := \#(I_z \cap I_i) + \#(I_c \cap I_i)$  and  $n_2 := \#(I_z \cap I_i)$ . In Problem (44), the following simplest cases can occur from the general ones described in Subsection 2.2:

- ad I.  $\mathbf{L} = \mathbf{2}$ :  $n_1 = 1, n_2 = 0$  (either one node in grazing contact with non-vanishing slip, or one node in strong contact with impending slip);
- ad IV.  $\mathbf{L} = \mathbf{4}, \mathbf{M}_1 = \mathbf{1}, \mathbf{M}_i = \mathbf{2}, i = \mathbf{2}, \mathbf{3}, \mathbf{4}$ :  $n_1 = 0, n_2 = 1$  (one node in grazing contact with impending slip);
- ad V.  $\mathbf{L} = \mathbf{4}, \mathbf{M}_i = \mathbf{2}, i = \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ :  $n_1 = 2, n_2 = 0$  (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 note that our abstract frame does not necessarily require nodal approximation of the contact conditions either. It covers also discrete problems arising from their integral approximation (combined with 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 (39),  $\alpha \in \{0, 1\}$ . The magnitude of  $|\tau - \tau_0|$  was decided to be large when it was greater than  $\tau_{\text{ref}}$ , and it was decided to be small when it was smaller than  $0.5\tau_{\text{ref}}$ .

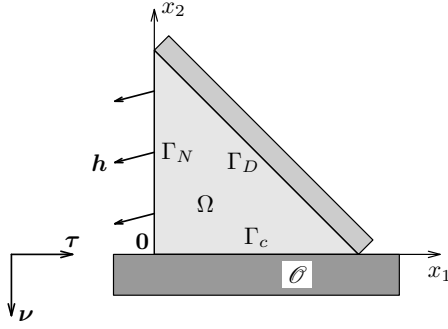


Fig. 13 Geometry of the problem with the triangular body.

## 5.1 Triangular Body

Let us consider contact of an isosceles triangle with a flat foundation (Fig. 13) in the framework of small-deformation elasticity with Lamé constants  $\lambda, \mu > 0$ . The triangle being fixed along  $\Gamma_D$  and volume forces being neglected, the model is parametrised via the surface force  $\mathbf{h} = \mathbf{h}(\gamma) = \gamma(h_1, h_2)$  with  $h_1$  and  $h_2$  constant. The friction coefficient  $\mathcal{F} > 0$  is supposed to be a fixed constant.

### 5.1.1 Discretisation with a Single Linear Element

First, we discretise the problem by using a single linear triangular finite element and prescribing the Dirichlet condition on  $\Gamma_D$  directly so that all degrees of freedom are related to the node in  $\mathbf{0}$ .

This discretisation leads to Problem (44) with  $\mathbf{H}: \mathbb{R}^5 \rightarrow \mathbb{R}^4$ ,

$$\mathbf{H}(\mathbf{y}) = \begin{pmatrix} au_v - bu_\tau + \gamma L_2 - \lambda_v \\ -bu_v + au_\tau - \gamma L_1 - \lambda_\tau \\ \lambda_v - (\lambda_v - u_v)_- \\ \lambda_\tau - P_{[\mathcal{F}(\lambda_v - u_v)_-, -\mathcal{F}(\lambda_v - u_v)_-]}(\lambda_\tau - u_\tau) \end{pmatrix},$$

$$\mathbf{y} := (\gamma, u_v, u_\tau, \lambda_v, \lambda_\tau) \in \mathbb{R}^5,$$

where  $a := (\lambda + 3\mu)/2$ ,  $b := (\lambda + \mu)/2$  and the load vector is given by  $\mathbf{L} = \mathbf{L}(\gamma) = \gamma(L_1, L_2)$ ,  $L_1$  and  $L_2$  being constant.

Let us take  $\bar{\mathbf{y}} := \mathbf{0}$ , 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,  $\mathbf{H} = \mathbf{H}^{(i)}$  in  $D^{(i)}$ ,  $i \in \mathcal{I}(\bar{\mathbf{y}}) = \{1, 2, 3, 4\}$ , with

$$D^{(1)} := \{\mathbf{y} \in \mathbb{R}^5; \lambda_v - u_v \geq 0\},$$

$$D^{(2)} := \{\mathbf{y} \in \mathbb{R}^5; \lambda_v - u_v \leq 0, \lambda_\tau - u_\tau \leq \mathcal{F}(\lambda_v - u_v)\},$$

$$D^{(3)} := \{\mathbf{y} \in \mathbb{R}^5; \mathcal{F}(\lambda_v - u_v) \leq \lambda_\tau - u_\tau \leq -\mathcal{F}(\lambda_v - u_v)\},$$

$$D^{(4)} := \{\mathbf{y} \in \mathbb{R}^5; \lambda_v - u_v \leq 0, -\mathcal{F}(\lambda_v - u_v) \leq \lambda_\tau - u_\tau\},$$

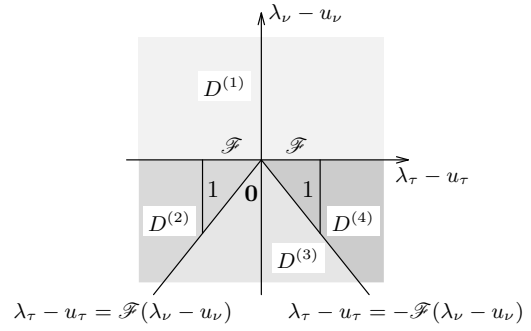


Fig. 14 Structure of the regions for the simple model.

$$\mathbf{H}^{(1)}(\mathbf{y}) := \begin{pmatrix} au_v - bu_\tau + \gamma L_2 - \lambda_v \\ -bu_v + au_\tau - \gamma L_1 - \lambda_\tau \\ \lambda_v \\ \lambda_\tau \end{pmatrix},$$

$$\mathbf{H}^{(2)}(\mathbf{y}) := \begin{pmatrix} au_v - bu_\tau + \gamma L_2 - \lambda_v \\ -bu_v + au_\tau - \gamma L_1 - \lambda_\tau \\ u_v \\ \mathcal{F}u_v - \mathcal{F}\lambda_v + \lambda_\tau \end{pmatrix},$$

$$\mathbf{H}^{(3)}(\mathbf{y}) := \begin{pmatrix} au_v - bu_\tau + \gamma L_2 - \lambda_v \\ -bu_v + au_\tau - \gamma L_1 - \lambda_\tau \\ u_v \\ u_\tau \end{pmatrix},$$

$$\mathbf{H}^{(4)}(\mathbf{y}) := \begin{pmatrix} au_v - bu_\tau + \gamma L_2 - \lambda_v \\ -bu_v + au_\tau - \gamma L_1 - \lambda_\tau \\ u_v \\ -\mathcal{F}u_v + \mathcal{F}\lambda_v + \lambda_\tau \end{pmatrix}$$

(Fig. 14). In this case, Problem ( $\mathcal{P}$ ) coincides with (NF) with  $\mathbf{A}^{(i)} = \nabla \mathbf{H}^{(i)}(\bar{\mathbf{y}})$  and  $C^{(i)} = D^{(i)}$ ,  $i \in \mathcal{I} = \{1, 2, 3, 4\}$ .

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

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

Further, Assumption 3 holds if

$$bL_1 - aL_2 \neq 0 \quad \text{and} \quad L_1 \pm \mathcal{F}L_2 \neq 0. \quad (46)$$

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

Clearly, the branching scenarios from  $\bar{\mathbf{y}}$  correspond to Case IV from Subsection 2.2. There are always at least two solution rays under satisfaction of (46), and one obtains by elementary calculations that there are even four solution rays if

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

Considering the case

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

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

$$\begin{aligned} \mathbf{y}^{(1)} &:= \left( 1, \frac{bL_1 - aL_2}{a^2 - b^2}, \frac{aL_1 - bL_2}{a^2 - b^2}, 0, 0 \right) && \text{(no contact),} \\ \mathbf{y}^{(2)} &:= \left( -1, 0, -\frac{L_1 + \mathcal{F}L_2}{a + b\mathcal{F}}, \frac{bL_1 - aL_2}{a + b\mathcal{F}}, \mathcal{F}\frac{bL_1 - aL_2}{a + b\mathcal{F}} \right) && \text{(contact-positive slip),} \\ \mathbf{y}^{(3)} &:= (1, 0, 0, L_2, -L_1) && \text{(contact-stick),} \\ \mathbf{y}^{(4)} &:= \left( 1, 0, \frac{L_1 - \mathcal{F}L_2}{a - b\mathcal{F}}, \frac{aL_2 - bL_1}{a - b\mathcal{F}}, \mathcal{F}\frac{bL_1 - aL_2}{a - b\mathcal{F}} \right) && \text{(contact-negative slip),} \end{aligned} \quad (47)$$

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

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

which implies that the solution ray generated by  $\mathbf{y}^{(1)}$  is coherently oriented with the ones generated by  $\mathbf{y}^{(2)}$  and  $\mathbf{y}^{(4)}$ , but incoherently oriented with the one generated by  $\mathbf{y}^{(3)}$ , 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  $\lambda = 100 \text{ GN/m}^2$ ,  $\mu = 82 \text{ GN/m}^2$ ,  $\mathbf{h}(\gamma) = \gamma(-26 \text{ GN/m}^2, -7.5 \text{ GN/m}^2)$  and  $\mathcal{F} = 1.7$ .

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 (44) and emanating from  $\bar{\mathbf{y}} := \mathbf{0}$ : one with  $\gamma$  negative and three with  $\gamma$  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. 15). According to the obvious correspondence of these branches to the ones from the simple model generated by  $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(4)}$  from (47), they are denoted as Branch 2, Branch 1, Branch 3 and Branch 4, respectively, here and in what follows. Nevertheless, observe that the non-uniqueness bound

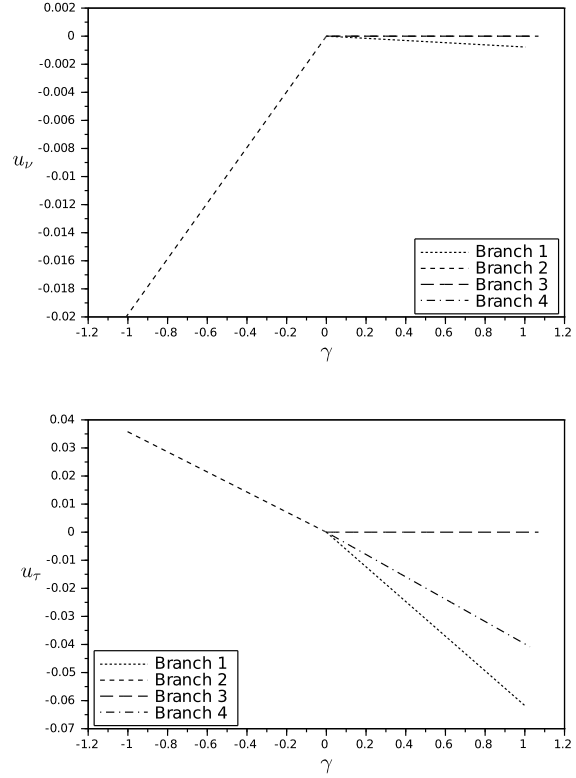


Fig. 16 Bifurcation diagrams.

for the friction coefficient is relaxed here:

$$\mathcal{F} = 1.7 < \frac{a}{b} = \frac{\lambda + 3\mu}{\lambda + \mu} \approx 1.9!$$

The bifurcation diagrams in Fig. 16 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 to the  $x_2$ -coordinate of the displacement) and to the normal contact stress at the nodes that are either in contact or not in contact with the foundation for two different branches simultaneously.

When  $\mathbf{y}_{k-1}$  and  $\mathbf{y}_k$  are chosen from various branches, computations by Algorithm 1 show that the numbers of the roots of  $\mathbf{J}(\alpha)$  defined by (37) in  $[0, 1]$  vary from zero to two: There is no root for  $\mathbf{y}_{k-1}, \mathbf{y}_k$  from the pairs  $\{\text{Branch 1, Branch 2}\}$  and  $\{\text{Branch 2, Branch 3}\}$ , one root for  $\{\text{Branch 1, Branch 3}\}$  and  $\{\text{Branch 2, Branch 4}\}$ , and two roots for  $\{\text{Branch 1, Branch 4}\}$  and  $\{\text{Branch 3, Branch 4}\}$ . 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 correspond exactly to the ones from the simple example although

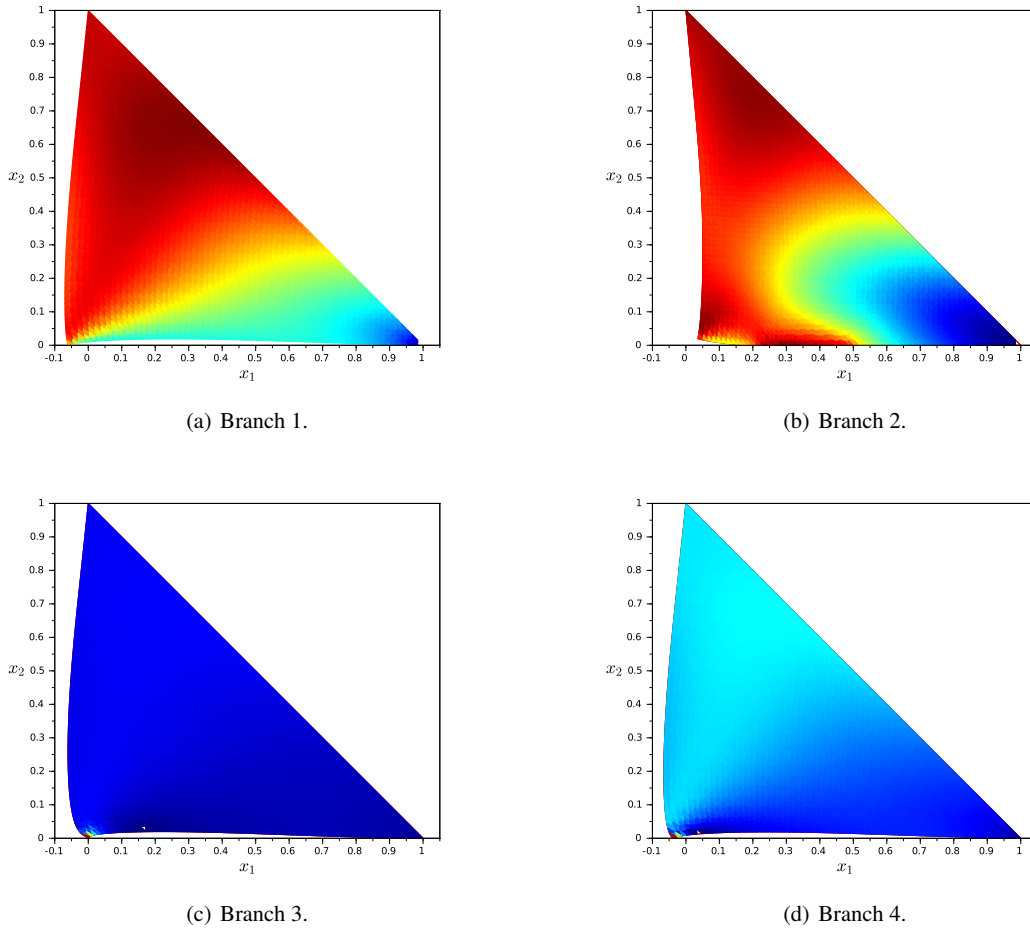


Fig. 15 Deformed bodies corresponding to the solutions with  $|\gamma| = 1$ : (b)  $\gamma = -1$ ; (a), (c), (d)  $\gamma = 1$ .

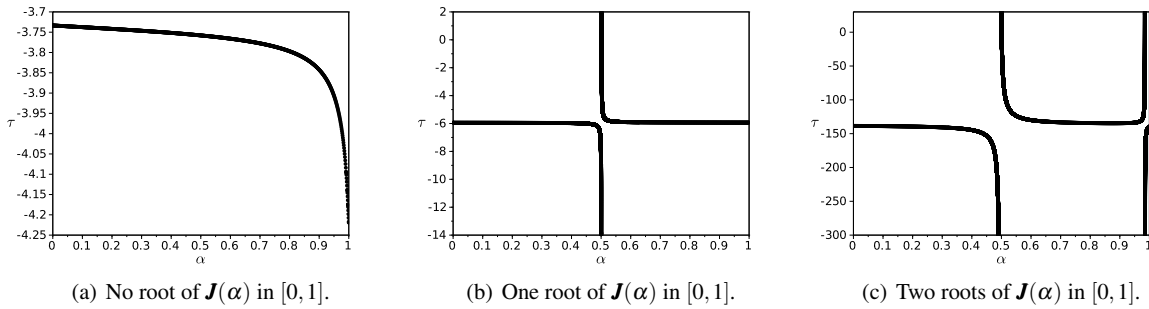


Fig. 17 Typical behaviours of  $\tau(\alpha)$ .

the overall situation is much more complex now – there are  $4^{64}$  regions intersecting at  $\bar{\mathbf{y}}$  in the present problem!

In the vast majority of our tests, there was no singularity of  $\tau(\alpha)$  introduced by (39) in  $[0, 1]$ , that is, no root of  $\mathbf{M}(\alpha)$  defined by (38) when there was no root of  $\mathbf{J}(\alpha)$  (see Fig. 17(a) for a typical behaviour of  $\tau(\alpha)$  in this case). Further, one and two roots of  $\mathbf{J}(\alpha)$  were usually closely

accompanied by one and two roots of  $\mathbf{M}(\alpha)$ , respectively (Figs. 17(b) and 17(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

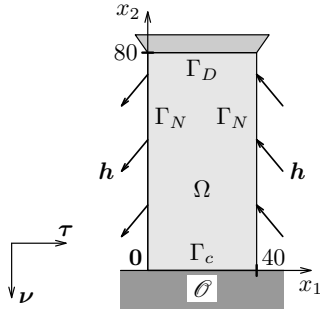


Fig. 18 Geometry of the problem with the rectangular body.

foundation, see Fig. 18. A plane-strain approximation of the nonlinear Ciarlet-Geymonat constitutive law [8, Chapter 4] is used:

$$\begin{aligned} \hat{\boldsymbol{\sigma}}(\mathbf{x}, \mathbf{F}) &= (\tilde{\boldsymbol{\sigma}}(\tilde{\mathbf{F}}))_{1 \leq i, j \leq 2}, \quad \tilde{\mathbf{F}} = \begin{pmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}, \quad \mathbf{F} \in \mathbb{R}^{2 \times 2}, \\ \tilde{\boldsymbol{\sigma}}(\tilde{\mathbf{F}}) &= 2b(\text{tr}(\tilde{\mathbf{F}}^\top \tilde{\mathbf{F}}))\mathbf{I} + 2(a - b\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\top)\tilde{\mathbf{F}} \\ &\quad + (2c\det(\tilde{\mathbf{F}}^\top \tilde{\mathbf{F}}) - d)\tilde{\mathbf{F}}^{-\top}, \quad \tilde{\mathbf{F}} \in \mathbb{R}^{3 \times 3}, \end{aligned}$$

where

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

and

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

The block is fixed along  $\Gamma_D$ , volume forces are neglected while the surface forces of the density given by the formula

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

(in  $\text{N/mm}^2$ ) act on both parts of  $\Gamma_N$  and  $\mathcal{F} = 1$  on  $\Gamma_c$ . The block is discretised by a uniform mesh with 800 bilinear squares and 21 contact nodes.

We have found numerically six solution branches emanating from  $\bar{\mathbf{y}} := \mathbf{0}$  in this problem: three with  $\gamma$  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  $\gamma$  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. 19).

The bifurcation diagrams in Fig. 20 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 branches can be continued by the same components of other branches without loss of differentiability, but the upper diagram clearly shows that this is not the case of all components.

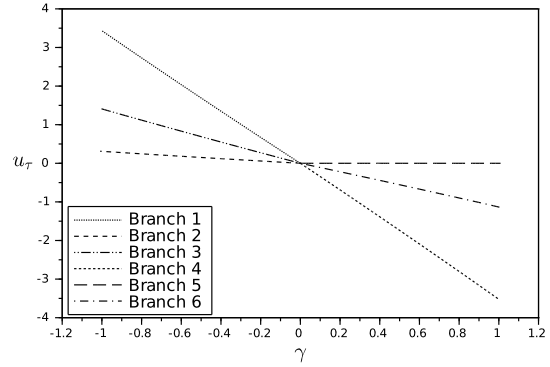
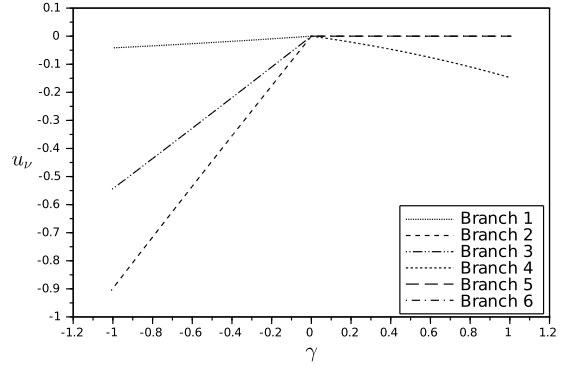
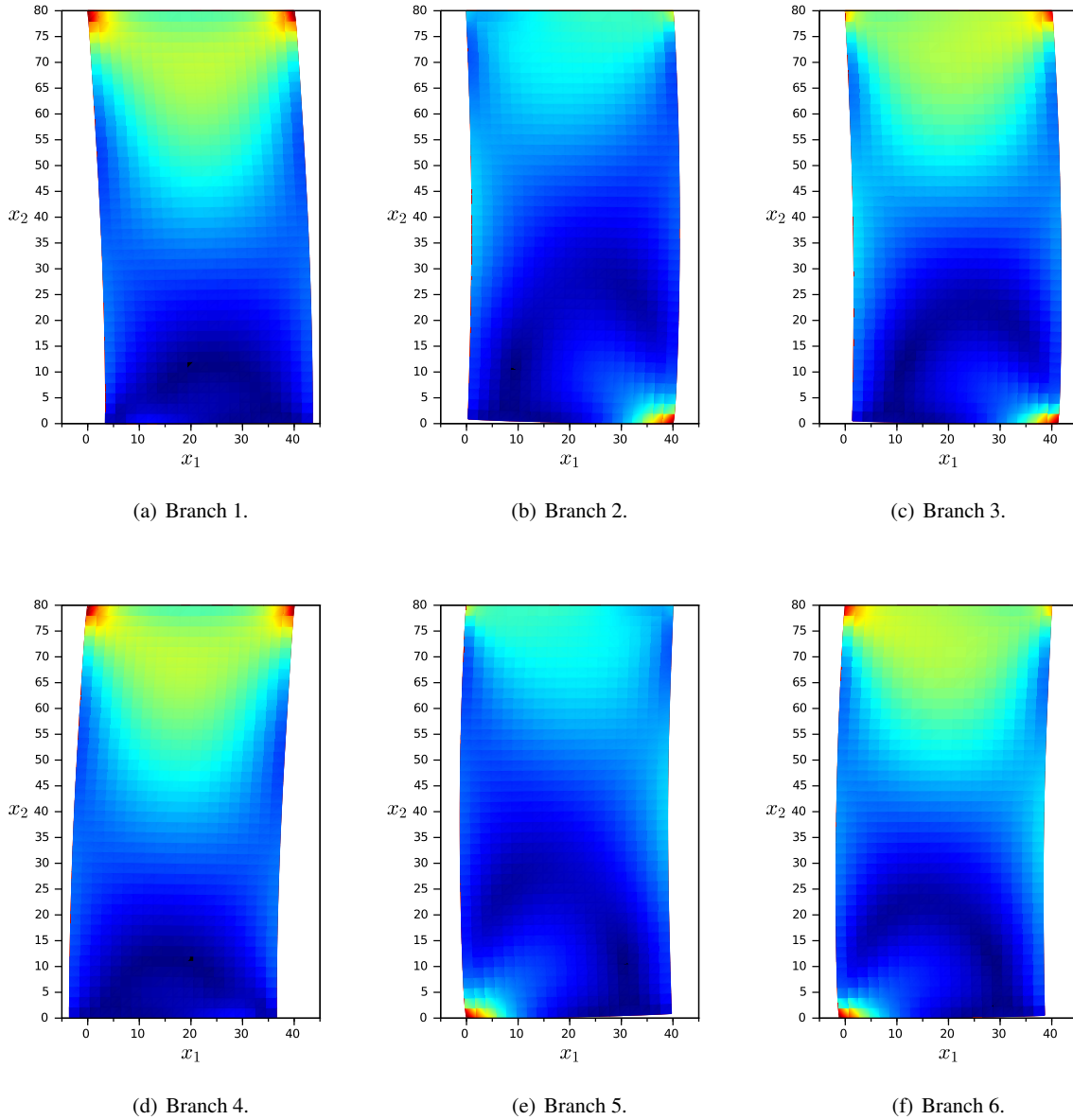


Fig. 20 Bifurcation diagrams.

The numbers of the roots of  $\mathbf{J}(\alpha)$  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  $\mathbf{J}(\alpha)$  do not differ significantly from Sub-section 5.1.2.

## 6 Conclusion

The paper presents a complex study of bifurcations for 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 (Theorem 2) completes the theoretical analysis of local behaviour of the solution set around a non-smooth point from [21]. In particular, the existence of solution curves is guaranteed in the directions determined by a simplified problem. It is worth mentioning that apart from being interesting for theoretical studies of branching via this simplified problem, our results can also be used for constructing methods of numerical continuation of the predictor-corrector type by providing (all possible) tangential predictions.



**Fig. 19** Deformed bodies corresponding to the solutions with  $|\gamma| = 1$ : (a), (b), (c)  $\gamma = -1$ ; (d), (e), (f)  $\gamma = 1$ .

**Table 1** The numbers of the roots of  $J(\alpha)$  in  $[0, 1]$ .

Branch	1	2	3	4	5	6
1	1	2	0	0	1	1
2	1	0	0	0	0	1
3	2	0	1	1	1	2
4	0	0	1	1	1	2
5	0	0	1	1	0	0
6	1	1	2	2	0	1

**Table 2** The orientations of the branches: the plus and minus signs stand for coherent and incoherent orientations, respectively.

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

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 piecewise-smooth function involved, and can thus be easily incorporated into 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.

We hope that our contribution illuminates the subject of piecewise-smooth bifurcations and sets up fundamentals for their numerical treatment.

Let us note that we have studied only a bifurcation problem of a given finite dimension, which corresponds typically to a given discretisation of a continuous problem. As observed in Subsection 5.1, the structure of the solution set may depend on the discretisation used. It would be a prospect of this work to investigate behaviour of the solution set with respect to various discretisations, especially, various discretisation parameters (mesh sizes).

**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). Besides, both authors would like to thank anonymous referees for careful reading of the manuscript and their stimulating comments that have helped to improve the exposition.

## References

- 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)
- Alexander, J.C.: The topological theory of an embedding method. In: H. Wacker (ed.) *Continuation methods*, pp. 37–68. Academic Press, New York (1978)
- 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)
- Allgower, E.L., Georg, K.: Numerical path following. In: *Handbook of numerical analysis*, vol. V, pp. 3–207. North-Holland, Amsterdam (1997)
- 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)
- 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)
- Björkman, G.: Path following and critical points for contact problems. *Comput. Mech.* **10**(3–4), 231–246 (1992)
- Ciarlet, P.G.: *Mathematical elasticity*. Vol. I: Three-dimensional elasticity, *Studies in Mathematics and its Applications*, vol. 20. North-Holland, Amsterdam (1988)
- Conrad, F., Herbin, R., Mittelmann, H.D.: Approximation of obstacle problems by continuation methods. *SIAM J. Numer. Anal.* **25**(6), 1409–1431 (1988)
- Conrad, F., Issard-Roch, F., Brauner, C.M., Nicolaenko, B.: Non-linear eigenvalue problems in elliptic variational inequalities: a local study. *Comm. Partial Differential Equations* **10**(2), 151–190 (1985)
- 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)
- 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)
- Decker, D.W., Keller, H.B.: Multiple limit point bifurcation. *J. Math. Anal. Appl.* **75**(2), 417–430 (1980)
- Facchinei, F., Pang, J.S.: *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Vol. I. Springer Series in Operations Research. Springer-Verlag, New York (2003)
- Georg, K.: Matrix-free numerical continuation and bifurcation. *Numer. Funct. Anal. Optim.* **22**(3–4), 303–320 (2001)
- 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)
- 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)
- Hild, P.: On finite element uniqueness studies for Coulomb’s frictional contact model. *Int. J. Appl. Math. Comput. Sci.* **12**(1), 41–50 (2002)
- 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)
- 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)
- 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)
- Lundberg, B.N., Poore, A.B.: Numerical continuation and singularity detection methods for parametric nonlinear programming. *SIAM J. Optim.* **3**(1), 134–154 (1993)
- Miersemann, E., Mittelmann, H.D.: Continuation for parametrized nonlinear variational inequalities. *J. Comput. Appl. Math.* **26**(1–2), 23–34 (1989)
- Paumier, J.C.: Bifurcations et méthodes numériques : applications aux problèmes elliptiques semi-linéaires, *Recherches en Mathématiques Appliquées*, vol. 42. Masson, Paris (1997)
- Poore, A.B., Tiaht, C.A.: Bifurcation problems in nonlinear parametric programming. *Math. Programming* **39**, 189–205 (1987)
- Renard, Y., Pommier, J.: GetFEM++. An open source generic C++ library for finite element methods, <http://home.gna.org/getfem>
- Scholtes, S.: Introduction to piecewise differentiable equations. Preprint No. 53/1994, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe (1994)
- Schulz, M., Pellegrino, S.: Equilibrium paths of mechanical systems with unilateral constraints I. Theory. *Proc. R. Soc. Lond. A* **456**(2001), 2223–2242 (2000)
- Tiaht, C.A., Poore, A.B.: A bifurcation analysis of the nonlinear parametric programming problem. *Math. Programming* **47**, 117–141 (1990)