



# On (local) analysis of multifunctions via subspaces contained in graphs of generalized derivatives



Helmut Gfrerer <sup>a,\*</sup>, Jiří V. Outrata <sup>b,c</sup>

<sup>a</sup> *Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040 Linz, Austria*

<sup>b</sup> *Institute of Information Theory and Automation, Czech Academy of Sciences, 18208 Prague, Czech Republic*

<sup>c</sup> *Centre for Informatics and Applied Optimization, Federation University of Australia, POB 663, Ballarat, Vic 3350, Australia*

## ARTICLE INFO

### Article history:

Received 1 June 2021

Available online 9 December 2021

Submitted by A. Daniilidis

### Keywords:

Generalized derivatives

Second-order theory

Strong metric (sub)regularity

Semismoothness\*

## ABSTRACT

The paper deals with a comprehensive theory of mappings, whose local behavior can be described by means of linear subspaces, contained in the graphs of two (primal and dual) generalized derivatives. This class of mappings includes the graphically Lipschitzian mappings and thus a number of multifunctions, frequently arising in optimization and equilibrium problems. The developed theory makes use of new generalized derivatives, provides us with some calculus rules and reveals a number of interesting connections. In particular, it enables us to construct a modification of the semismooth\* Newton method with improved convergence properties and to derive a generalization of Clarke's Inverse Function Theorem to multifunctions together with new efficient characterizations of strong metric (sub)regularity and tilt stability.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

When implementing the semismooth\* Newton method [14] for solving an inclusion of the form

$$0 \in F(x)$$

with some set-valued mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ , we observed that it is advantageous to work with linear subspaces  $L \subseteq \mathbb{R}^n \times \mathbb{R}^n$  having dimension  $n$  and contained in the graph of the limiting coderivative, i.e.,

$$L \subseteq \text{gph } D^*F(x, y)$$

at points  $(x, y) \in \text{gph } F$ .

\* Corresponding author.

E-mail addresses: [helmut.gfrerer@jku.at](mailto:helmut.gfrerer@jku.at) (H. Gfrerer), [outrata@utia.cas.cz](mailto:outrata@utia.cas.cz) (J.V. Outrata).

However, this paper goes far beyond the analysis of the above issue and presents a comprehensive study of a class of mappings, whose local behavior can be described by appropriately constructed linear subspaces. This class turns out to be rather broad and the developed theory helps us both to suggest an efficient modification of the semismooth\* Newton method as well as to derive a number of new results concerning strong metric subregularity, strong metric regularity and tilt stability. More precisely, for the mentioned mapping  $F$ , primal and dual generalized derivatives are introduced, whose elements are subspaces of dimension  $n$ . In order to define these derivatives, we consider points in the graph of the mapping where the tangent cone amounts to a subspace and then perform an outer limiting operation in a certain compact metric space.

Our construction is motivated by the definition of the B-subdifferential (Bouligand-subdifferential) for single-valued mappings, whose elements are given as limit of Jacobians at points where the mapping is Fréchet differentiable. Note that the tangent cone to the graph of a map is a subspace whenever the mapping is differentiable at the point under consideration. Instead of computing limits of matrices, we consider the limit of subspaces given by the graph of the linear mappings induced by the matrices. When the mapping is Lipschitzian, then we obtain a one-to-one correspondence between the new primal generalized derivative and the B-subdifferential. However, for non-Lipschitzian single-valued mappings there will be a difference because we are considering limits of subspaces in a compact metric space whereas the underlying matrices can be unbounded.

There are some relations between our generalized derivatives and existing ones. The dual derivative consists of subspaces contained in the limiting coderivative and the elements of the primal derivative are subspaces contained in the so-called outer limiting graphical derivative. To the best of our knowledge, the latter has not yet been considered in the literature and is contained in the so-called strict graphical derivative.

Our theory is not applicable to arbitrary mappings. However, as already mentioned, the class of mappings which are suited for our approach, is rather broad and important for applications. In particular, every mapping which is graphically Lipschitzian, i.e., its graph coincides under some change of coordinates with the graph of a locally Lipschitzian mapping, belongs to this class. Graphically Lipschitzian mappings have been already considered by Rockafellar [33]. E.g., locally maximally hypomonotone mappings like the subdifferential mapping of prox-regular and subdifferentially continuous functions possess this property [27]. Thus, our approach is particularly suitable for second-order theory and we will establish a strong relationship with the so-called quadratic bundle introduced in the recent paper [35]. Note that in [33] also a limit of tangent spaces has been considered. However, in [33] an inner limit with respect to the usual set-convergence has been used yielding a different sort of results.

Within the framework of the new theory one can introduce a new regularity notion leading to an adaptation of the semismooth\* Newton method. This notion is weaker than metric regularity and enables us to streamline the algorithm and to relax the assumptions, ensuring its locally superlinear convergence. Under the respective regularity condition it is also possible to show that a semismooth\* mapping is strongly metrically subregular, not only at the reference point itself but also on a neighborhood of it. It seems that this somewhat extended property of strong metric subregularity *around* the reference point has not been considered yet. As a byproduct, we present a characterization of this property by means of the outer limiting graphical derivative.

Finally we turn our attention to the property of strong metric regularity. Since strongly regular mappings are graphically Lipschitzian by the definition, the preceding theory enables us to reveal some interesting new connections. In particular, one obtains a generalization of Clarke's Inverse Function Theorem to set-valued mappings and, when applied to specific problem classes, these results lead to new characterizations of strong metric regularity for locally maximally monotone operators and to a new characterization of tilt-stability. Compared with existing characterizations, the new ones have the advantage, that not the whole strict graphical derivative or limiting coderivative must be checked (as, e.g., in [20,8,28]) but only a condition on

the subspaces contained in its graph. In this way the arsenal of available criteria for strong metric regularity (cf. [31,9,20,15,8,10]) and tilt stability (cf. [28,11,24]) is enriched.

The plan of the paper is as follows. After the Preliminaries, devoted to relevant notions from variational analysis, in Section 3 we introduce and analyze the crucial class of SCD (subspace containing derivative) mappings. In their analysis we make use of the mentioned generalized derivatives, for which some basic calculus rules are developed and exact formulas in case of graphically Lipschitzian mappings are provided. In Section 4 we introduce the notion of SCD-regularity, which plays a central role in the subsequent sections. Section 5 deals with the adaptation of the semismooth\* Newton method to SCD mappings. The property of strong metric subregularity around a point is characterized in Section 6. Finally, in Section 7 we present a generalization of Clarke’s Inverse Function Theorem and new characterizations of strong metric regularity and tilt stability for various classes of SCD mappings.

The following notation is employed. Given a linear subspace  $L \subseteq \mathbb{R}^n$ ,  $L^\perp$  denotes its orthogonal complement and, for a closed cone  $K$  with vertex at the origin,  $K^\circ$  signifies its (negative) polar. Further, given a multifunction  $F$ ,  $\text{gph } F := \{(x, y) \mid y \in F(x)\}$  stands for its graph. For an element  $u \in \mathbb{R}^n$ ,  $\|u\|$  denotes its Euclidean norm and  $\mathcal{B}_\delta(u)$  denotes the closed ball around  $u$  with radius  $\delta$ . In a product space we use the norm  $\|(u, v)\| := \sqrt{\|u\|^2 + \|v\|^2}$ . Given an  $m \times n$  matrix  $A$ , we employ the operator norm  $\|A\|$  with respect to the Euclidean norm and we denote the range of  $A$  by  $\text{rge } A$ . Given a set  $\Omega \subset \mathbb{R}^s$ , we define the distance of a point  $x$  to  $\Omega$  by  $d_\Omega(x) := \text{dist}(x, \Omega) := \inf\{\|y - x\| \mid y \in \Omega\}$  and the indicator function is denoted by  $\delta_\Omega$ . When a mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is differentiable at  $x$ , we denote by  $\nabla F(x)$  its Jacobian.

## 2. Preliminaries

Throughout the whole paper, we will frequently use the following basic notions of modern variational analysis. All the sets under consideration are supposed to be *locally closed* around the points in question without further mentioning.

**Definition 2.1.** Let  $A$  be a set in  $\mathbb{R}^s$  and let  $\bar{x} \in A$ . Then

- (i) The *tangent (contingent, Bouligand) cone* to  $A$  at  $\bar{x}$  is given by

$$T_A(\bar{x}) := \text{Lim sup}_{t \downarrow 0} \frac{A - \bar{x}}{t}$$

and the *paratingent cone* to  $A$  at  $\bar{x}$  is given by

$$T_A^P(\bar{x}) := \text{Lim sup}_{\substack{x \xrightarrow{A} \bar{x} \\ t \downarrow 0}} \frac{A - x}{t}$$

- (ii) The set

$$\widehat{N}_A(\bar{x}) := (T_A(\bar{x}))^\circ$$

is the *regular (Fréchet) normal cone* to  $A$  at  $\bar{x}$ , and

$$N_A(\bar{x}) := \text{Lim sup}_{x \xrightarrow{A} \bar{x}} \widehat{N}_A(x)$$

is the *limiting (Mordukhovich) normal cone* to  $A$  at  $\bar{x}$ . Given a direction  $d \in \mathbb{R}^s$ ,

$$N_A(\bar{x}; d) := \operatorname{Lim\,sup}_{d' \xrightarrow{t \downarrow 0} d} \widehat{N}_A(\bar{x} + td')$$

is the *directional limiting normal cone* to  $A$  at  $\bar{x}$  in direction  $d$ .

In this definition “Lim sup” stands for the Painlevé-Kuratowski *outer (upper) set limit*, see, e.g., [1]. If  $A$  is convex, then  $\widehat{N}_A(\bar{x}) = N_A(\bar{x})$  amounts to the classical normal cone in the sense of convex analysis and we will write  $N_A(\bar{x})$ . By the definition, the limiting normal cone coincides with the directional limiting normal cone in direction 0, i.e.,  $N_A(\bar{x}) = N_A(\bar{x}; 0)$ , and  $N_A(\bar{x}; d) = \emptyset$  whenever  $d \notin T_A(\bar{x})$ .

The above listed cones enable us to describe the local behavior of set-valued maps via various generalized derivatives. All the set-valued mappings under consideration are supposed to have *locally closed graph* around the points in question.

**Definition 2.2.** Consider a multifunction  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  and let  $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ .

- (i) The multifunction  $DF(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  given by  $\operatorname{gph} DF(\bar{x}, \bar{y}) = T_{\operatorname{gph} F}(\bar{x}, \bar{y})$  is called the *graphical derivative* of  $F$  at  $(\bar{x}, \bar{y})$ .
- (ii) The multifunction  $D_*F(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  given by  $\operatorname{gph} D_*F(\bar{x}, \bar{y}) = T_{\operatorname{gph} F}^P(\bar{x}, \bar{y})$  is called the *strict (paratingent) derivative* of  $F$  at  $(\bar{x}, \bar{y})$ .
- (iii) The multifunction  $\widehat{D}^*F(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  defined by

$$\operatorname{gph} \widehat{D}^*F(\bar{x}, \bar{y}) = \{(y^*, x^*) \mid (x^*, -y^*) \in \widehat{N}_{\operatorname{gph} F}(\bar{x}, \bar{y})\}$$

is called the *regular (Fréchet) coderivative* of  $F$  at  $(\bar{x}, \bar{y})$ .

- (iv) The multifunction  $D^*F(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ , defined by

$$\operatorname{gph} D^*F(\bar{x}, \bar{y}) = \{(y^*, x^*) \mid (x^*, -y^*) \in N_{\operatorname{gph} F}(\bar{x}, \bar{y})\}$$

is called the *limiting (Mordukhovich) coderivative* of  $F$  at  $(\bar{x}, \bar{y})$ .

- (v) Given a pair of directions  $(u, v) \in \mathbb{R}^n \times \mathbb{R}^m$ , the multifunction  $D^*F((\bar{x}, \bar{y}); (u, v)) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ , defined by

$$\operatorname{gph} D^*F((\bar{x}, \bar{y}); (u, v)) = \{(y^*, x^*) \mid (x^*, -y^*) \in N_{\operatorname{gph} F}((\bar{x}, \bar{y}); (u, v))\}$$

is called the *directional limiting coderivative* of  $F$  at  $(\bar{x}, \bar{y})$  in direction  $(u, v)$ .

The directional limiting normal cone and coderivative were introduced by the first author in [12] and various properties of these objects can be found also in [13] and in the references therein. Note that  $D^*F(\bar{x}, \bar{y}) = D^*F((\bar{x}, \bar{y}); (0, 0))$  and that  $\operatorname{dom} D^*F((\bar{x}, \bar{y}); (u, v)) = \emptyset$  whenever  $v \notin DF(\bar{x}, \bar{y})(u)$ .

Note that by [34, Proposition 6.6] and the definition of the limiting coderivative we have

$$\operatorname{gph} D^*F(\bar{x}, \bar{y}) = \operatorname{Lim\,sup}_{(x, y) \xrightarrow{\operatorname{gph} F} (\bar{x}, \bar{y})} \operatorname{gph} D^*F(x, y). \quad (1)$$

If  $F$  is single-valued, we can omit the second argument and write  $DF(x)$ ,  $\widehat{D}^*F(x)$ , ... instead of  $DF(x, F(x))$ ,  $\widehat{D}^*F(x, F(x))$ , ... However, be aware that when considering limiting objects at  $x$  where  $F$  is not continuous, it is not enough to consider only sequences  $x_k \rightarrow x$  but we must work with sequences  $(x_k, F(x_k)) \rightarrow (x, F(x))$ .

**Definition 2.3.** Let  $U \subset \mathbb{R}^n$  be open and let  $F : U \rightarrow \mathbb{R}^m$  be a mapping. The *B-subdifferential* of  $F$  at  $x \in U$  is defined as

$$\overline{\nabla}F(x) := \{A \mid \exists x_k \rightarrow x : F \text{ is Fréchet differentiable at } x_k \text{ and } A = \lim_{k \rightarrow \infty} \nabla F(x_k)\} \tag{2}$$

Recall that the Clarke Generalized Jacobian is given by  $\text{conv } \overline{\nabla}F(x)$ , i.e., the convex hull of the B-subdifferential.

There exists the following relation between the B-subdifferential and the coderivative of  $F$ , which states that every element from the B-subdifferential defines a certain subspace contained in the graph of the coderivative.

**Proposition 2.4.** Let  $U \subset \mathbb{R}^n$  be open and let  $F : U \rightarrow \mathbb{R}^m$  be a mapping. Let  $F$  be continuous at  $x \in U$  and let  $A \in \overline{\nabla}F(x)$ . Then

$$(y^*, A^T y^*) \in \text{gph } D^*F(x) \quad \forall y^* \in \mathbb{R}^m.$$

**Proof.** Consider  $A \in \overline{\nabla}F(x)$  together with some sequence  $x_k \rightarrow x$  such that  $\nabla F(x_k) \rightarrow A$  as  $k \rightarrow \infty$ . By [34, Example 9.25(b)] we have  $T_{\text{gph } F}(x_k, F(x_k)) = \{(u, \nabla F(x_k)u) \mid u \in \mathbb{R}^n\}$  and therefore  $(y^*, \nabla F(x_k)^T y^*) \in \text{gph } \hat{D}^*F(x_k), \forall y^* \in \mathbb{R}^m$ . By passing to the limit, the assertion follows from the definition of the limiting coderivative.  $\square$

If the mapping  $F : U \rightarrow \mathbb{R}^m$  is Lipschitz continuous, by Rademacher’s Theorem  $F$  is differentiable almost everywhere in  $U$  and  $\|\nabla F(x)\|$  is bounded there by the Lipschitz constant of  $F$ . Thus  $\overline{\nabla}F(\bar{x}) \neq \emptyset$  for Lipschitz continuous mappings  $F$ .

Let  $q : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  be an extended-real-valued function with the domain and the epigraph

$$\text{dom } q := \{x \in \mathbb{R}^n \mid q(x) < \infty\}, \quad \text{epi } q := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid \alpha \geq q(x)\}.$$

The (*limiting/Mordukhovich*) *subdifferential* of  $q$  at  $\bar{x} \in \text{dom } q$  is defined geometrically by

$$\partial q(\bar{x}) := \{x^* \in \mathbb{R}^n \mid (x^*, -1) \in N_{\text{epi } q}(\bar{x}, q(\bar{x}))\}.$$

This subdifferential is a general extension of the classical gradient for smooth functions and of the classical subdifferential of convex ones.

If  $q(\bar{x})$  is finite, define the parametric family of second-order difference quotients for  $q$  at  $\bar{x}$  for  $\bar{x}^* \in \mathbb{R}^n$  by

$$\Delta_t^2 q(\bar{x}, \bar{x}^*)(w) := \frac{q(\bar{x} + tw) - q(\bar{x}) - t\langle \bar{x}^*, w \rangle}{\frac{1}{2}t^2} \quad \text{with } w \in \mathbb{R}^n, t > 0.$$

The *second-order subderivative* of  $q$  at  $\bar{x}$  for  $\bar{x}^*$  is given by

$$d^2 q(\bar{x}, \bar{x}^*)(w) = \liminf_{\substack{t \downarrow 0 \\ w' \rightarrow w}} \Delta_t^2 q(\bar{x}, \bar{x}^*)(w').$$

$q$  is called *twice epi-differentiable* at  $\bar{x}$  for  $\bar{x}^*$ , if the functions  $\Delta_t^2 q(\bar{x}, \bar{x}^*)$  epi-converge to  $d^2 q(\bar{x}, \bar{x}^*)$  as  $t \downarrow 0$ .

Let us now recall the following regularity notions.

**Definition 2.5.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  be a mapping and let  $(\bar{x}, \bar{y}) \in \text{gph } F$ .

1.  $F$  is said to be *metrically subregular at*  $(\bar{x}, \bar{y})$  if there exists  $\kappa \geq 0$  along with some neighborhood  $X$  of  $\bar{x}$  such that

$$\text{dist}(x, F^{-1}(\bar{y})) \leq \kappa \text{dist}(\bar{y}, F(x)) \quad \forall x \in X. \quad (3)$$

The infimum over all  $\kappa \geq 0$  such that (3) holds for some neighborhood  $X$  is denoted by  $\text{subreg } F(\bar{x}, \bar{y})$ .

2.  $F$  is said to be *strongly metrically subregular at*  $(\bar{x}, \bar{y})$  if it is metrically subregular at  $(\bar{x}, \bar{y})$  and there exists a neighborhood  $X'$  of  $\bar{x}$  such that  $F^{-1}(\bar{y}) \cap X' = \{\bar{x}\}$ .
3.  $F$  is said to be *metrically regular around*  $(\bar{x}, \bar{y})$  if there is  $\kappa \geq 0$  together with neighborhoods  $X$  of  $\bar{x}$  and  $Y$  of  $\bar{y}$  such that

$$\text{dist}(x, F^{-1}(y)) \leq \kappa \text{dist}(y, F(x)) \quad \forall (x, y) \in X \times Y. \quad (4)$$

The infimum over all  $\kappa \geq 0$  such that (4) holds for some neighborhoods  $X, Y$  is denoted by  $\text{reg } F(\bar{x}, \bar{y})$ .

4.  $F$  is said to be *strongly metrically regular around*  $(\bar{x}, \bar{y})$  if it is metrically regular around  $(\bar{x}, \bar{y})$  and  $F^{-1}$  has a single-valued localization around  $(\bar{y}, \bar{x})$ , i.e., there are open neighborhoods  $Y'$  of  $\bar{y}$ ,  $X'$  of  $\bar{x}$  and a mapping  $h : Y' \rightarrow \mathbb{R}^n$  with  $h(\bar{y}) = \bar{x}$  such that  $\text{gph } F \cap (X' \times Y') = \{(h(y), y) \mid y \in Y'\}$ .

It is well-known, see, e.g., [10] that the property of (strong) metric subregularity for  $F$  at  $(\bar{x}, \bar{y})$  is equivalent with the property of (*isolated*) *calmness* for  $F^{-1}$  at  $(\bar{y}, \bar{x})$ . Further,  $F$  is metrically regular around  $(\bar{x}, \bar{y})$  if and only if the inverse mapping  $F^{-1}$  has the so-called *Aubin property* around  $(\bar{y}, \bar{x})$ . In this paper we will frequently use the following characterization of strong metric regularity.

**Theorem 2.6** (cf. [10, Proposition 3G.1]).  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  if and only if  $F^{-1}$  has a Lipschitz continuous localization  $h$  around  $(\bar{y}, \bar{x})$ . In this case there holds

$$\text{reg } F(\bar{x}, \bar{y}) = \limsup_{\substack{y, y' \rightarrow \bar{y} \\ y \neq y'}} \frac{\|h(y) - h(y')\|}{\|y - y'\|}.$$

In this paper we will also use the following point-based characterizations of the above regularity properties.

**Theorem 2.7.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  be a mapping and let  $(\bar{x}, \bar{y}) \in \text{gph } F$ .

- (i) (*Levy-Rockafellar criterion*)  $F$  is strongly metrically subregular at  $(\bar{x}, \bar{y})$  if and only if

$$0 \in DF(\bar{x}, \bar{y})(u) \Rightarrow u = 0, \quad (5)$$

and in this case one has

$$\text{subreg } F(\bar{x}, \bar{y}) = \sup\{\|u\| \mid (u, v) \in \text{gph } DF(\bar{x}, \bar{y}), \|v\| \leq 1\}.$$

- (ii) (*Mordukhovich criterion*)  $F$  is metrically regular around  $(\bar{x}, \bar{y})$  if and only if

$$0 \in D^*F(\bar{x}, \bar{y})(y^*) \Rightarrow y^* = 0. \quad (6)$$

Further, in this case one has

$$\text{reg } F(\bar{x}, \bar{y}) = \sup\{\|y^*\| \mid (y^*, x^*) \in \text{gph } D^*F(\bar{x}, \bar{y}), \|x^*\| \leq 1\}. \quad (7)$$

(iii)  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  if and only if

$$0 \in D_*F(\bar{x}, \bar{y})(u) \Rightarrow u = 0 \tag{8}$$

and (6) holds. In this case one also has

$$\text{reg } F(\bar{x}, \bar{y}) = \sup\{\|u\| \mid (u, v) \in \text{gph } D_*F(\bar{x}, \bar{y}), \|v\| \leq 1\}. \tag{9}$$

**Proof.** Statement (i) follows from [10, Theorem 4E.1], see also [21]. Statement (ii) can be found in [22, Theorem 3.3]. The criterion for strong metric regularity follows from Dontchev and Frankowska [8, Theorem 16.2] by taking into account that the condition  $\bar{x} \in \liminf_{y \rightarrow \bar{y}} F^{-1}(y)$  appearing in [8, Theorem 16.2] can be ensured by the requirement that  $F$  is metrically regular which in turn can be characterized by the Mordukhovich criterion.  $\square$

For a sufficient condition for metric subregularity based on directional limiting coderivatives we refer to [13].

The properties of (strong) metric regularity and strong metric subregularity are stable under Lipschitzian and calm perturbations, respectively, cf. [10]. Further note that the property of (strong) metric regularity holds around all points belonging to the graph of  $F$  sufficiently close to the reference point, whereas the property of (strong) metric subregularity is guaranteed to hold only at the reference point. This leads to the following definition.

**Definition 2.8.** We say that the mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is (strongly) metrically subregular around  $(\bar{x}, \bar{y}) \in \text{gph } F$  if there is a neighborhood  $W$  of  $(\bar{x}, \bar{y})$  such that  $F$  is (strongly) metrically subregular at every point  $(x, y) \in \text{gph } F \cap W$  and we define

$$\text{l-subreg } F(\bar{x}, \bar{y}) := \limsup_{(x,y) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})} \text{subreg } F(x, y) < \infty.$$

In this case we will also speak about (strong) metric subregularity on a neighborhood.

Note that every polyhedral multifunction, i.e., a mapping whose graph is the union of finitely many convex polyhedral sets, is metrically subregular around every point of its graph by Robinson’s result [32]. In Section 6, characterizations of strong metric subregularity on a neighborhood will be investigated.

Next we introduce the semismooth\* sets and mappings.

**Definition 2.9** (cf. [14]).

1. A set  $A \subseteq \mathbb{R}^s$  is called *semismooth\** at a point  $\bar{x} \in A$  if for all  $u \in \mathbb{R}^s$  it holds

$$\langle x^*, u \rangle = 0 \quad \forall x^* \in N_A(\bar{x}; u). \tag{10}$$

2. A set-valued mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is called *semismooth\** at a point  $(\bar{x}, \bar{y}) \in \text{gph } F$ , if  $\text{gph } F$  is semismooth\* at  $(\bar{x}, \bar{y})$ , i.e., for all  $(u, v) \in \mathbb{R}^n \times \mathbb{R}^m$  we have

$$\langle u^*, u \rangle = \langle v^*, v \rangle \quad \forall (v^*, u^*) \in \text{gph } D^*F((\bar{x}, \bar{y}); (u, v)). \tag{11}$$

The class of semismooth\* mappings is rather broad. We list here two important classes of multifunctions having this property.

**Proposition 2.10.**

- (i) Every mapping whose graph is the union of finitely many closed convex sets is semismooth\* at every point of its graph.
- (ii) Every mapping with closed subanalytic graph is semismooth\* at every point of its graph.

**Proof.** The first assertion was already shown in [14, Proposition 3.4, 3.5]. As mentioned in [14, Remark 3.10], the semismooth\* property of sets amounts to the notion of semismoothness introduced in [16]. It follows thus from [19, Theorem 2], that all closed subanalytic sets are automatically semismooth\* and the second statement holds by the definition of semismooth\* mappings.  $\square$

The statement of Proposition 2.10(ii) can be considered as the counterpart to [4], where it is shown that locally Lipschitz tame mappings  $F : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  are semismooth in the sense of Qi and Sun [30]. In case of single-valued Lipschitzian mappings the semismooth\* property is equivalent with the semismooth property introduced by Gowda [15], which is weaker than the one in [30].

In the above definition the semismooth\* sets and mappings have been defined via directional limiting normal cones and coderivatives. For our purpose it is convenient to make use of equivalent characterizations in terms of standard (regular and limiting) normal cones and coderivatives, respectively.

**Proposition 2.11** (cf. [14, Corollary 3.3]). *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  and  $(\bar{x}, \bar{y}) \in \text{gph } F$  be given. Then the following three statements are equivalent*

- (i)  $F$  is semismooth\* at  $(\bar{x}, \bar{y})$ .
- (ii) For every  $\varepsilon > 0$  there is some  $\delta > 0$  such that

$$\begin{aligned} |\langle x^*, x - \bar{x} \rangle - \langle y^*, y - \bar{y} \rangle| &\leq \varepsilon \|(x, y) - (\bar{x}, \bar{y})\| \|(x^*, y^*)\| \\ \forall (x, y) \in \mathcal{B}_\delta(\bar{x}, \bar{y}) \quad \forall (y^*, x^*) \in \text{gph } \widehat{D}^*F(x, y). \end{aligned} \quad (12)$$

- (iii) For every  $\varepsilon > 0$  there is some  $\delta > 0$  such that

$$\begin{aligned} |\langle x^*, x - \bar{x} \rangle - \langle y^*, y - \bar{y} \rangle| &\leq \varepsilon \|(x, y) - (\bar{x}, \bar{y})\| \|(x^*, y^*)\| \\ \forall (x, y) \in \mathcal{B}_\delta(\bar{x}, \bar{y}) \quad \forall (y^*, x^*) \in \text{gph } D^*F(x, y). \end{aligned} \quad (13)$$

**3. SCD mappings**

In what follows we denote by  $\mathcal{L}_n$  the metric space of all  $n$ -dimensional subspaces of  $\mathbb{R}^{2n}$  equipped with the metric

$$d_{\mathcal{L}}(L_1, L_2) := \|P_1 - P_2\|$$

where  $P_i$  is the symmetric  $2n \times 2n$  matrix representing the orthogonal projection on  $L_i$ ,  $i = 1, 2$ .

Sometimes we will also work with bases for the subspaces  $L \in \mathcal{L}_n$ . Let  $\mathcal{M}_n$  denote the collection of all  $2n \times n$  matrices with full rank  $n$  and for  $L \in \mathcal{L}_n$  we define

$$\mathcal{M}(L) := \{Z \in \mathcal{M}_n \mid \text{rge } Z = L\},$$

i.e., the columns of  $Z \in \mathcal{M}(L)$  are a basis for  $L$ . Further we denote by  $\mathcal{M}^{\text{orth}}(L)$  the set of all matrices  $Z \in \mathcal{M}(L)$  with  $Z^T Z = I$ , i.e., the columns of  $Z$  are an orthogonal basis for  $L$ . Recall that, given any matrix  $\bar{Z} \in \mathcal{M}(L)$  (or  $\bar{Z} \in \mathcal{M}^{\text{orth}}(L)$ ), there holds



$$\mathcal{M}(L) = \{\bar{Z}B \mid B \text{ nonsingular } n \times n \text{ matrix}\} \quad (\mathcal{M}^{\text{orth}}(L) = \{\bar{Z}B \mid B \text{ orthogonal } n \times n \text{ matrix}\}).$$

Further recall that the  $2n \times 2n$  matrix  $P$ , representing the orthogonal projection on some  $L \in \mathcal{L}_n$ , admits the representations

$$P = Z(Z^T Z)^{-1} Z^T, \quad Z \in \mathcal{M}(L) \quad \text{and} \quad P = ZZ^T, \quad Z \in \mathcal{M}^{\text{orth}}(L). \tag{14}$$

**Lemma 3.1.**

- (i) Let  $Z_k \in \mathcal{M}_n$  be a sequence converging to some  $Z \in \mathcal{M}_n$ . Then  $\text{rge } Z_k$  converges in  $\mathcal{L}_n$  to  $\text{rge } Z \in \mathcal{L}_n$ .
- (ii) Let  $L_k \in \mathcal{L}_n$  be a sequence converging to  $L \in \mathcal{L}_n$ . Then there is a sequence  $Z_k \in \mathcal{M}(L_k)$  converging to some  $Z \in \mathcal{M}(L)$ .
- (iii) Let  $A_k$  be a sequence of nonsingular  $2n \times 2n$  matrices converging to a nonsingular matrix  $A$  and let  $L_k \in \mathcal{L}_n$  be a sequence converging to  $L \in \mathcal{L}_n$ . Then  $\lim_{k \rightarrow \infty} d_{\mathcal{L}}(A_k L_k, AL) = 0$ .
- (iv) The metric space  $\mathcal{L}_n$  is compact.
- (v) Let  $L_k \in \mathcal{L}_n$  be a sequence and let  $L \in \mathcal{L}_n$ . Then  $L_k$  converges to  $L$  in  $\mathcal{L}_n$ , i.e.,  $\lim_{k \rightarrow \infty} d_{\mathcal{L}}(L_k, L) = 0$ , if and only if  $\lim_{k \rightarrow \infty} L_k = L$  in the sense of Painlevé-Kuratowski convergence.

**Proof.** The first statement follows immediately from  $Z(Z^T Z)^{-1} Z^T = \lim_{k \rightarrow \infty} Z_k(Z_k^T Z_k)^{-1} Z_k^T$  together with (14). In order to prove (ii), choose  $Z_k \in \mathcal{M}^{\text{orth}}(L_k)$  and  $Z \in \mathcal{M}^{\text{orth}}(L)$ . Then  $ZZ^T = \lim_{k \rightarrow \infty} Z_k Z_k^T$  due to  $L_k \xrightarrow{\mathcal{L}_n} L$  and consequently  $Z = ZZ^T Z = \lim_{k \rightarrow \infty} \tilde{Z}_k$  with  $\tilde{Z}_k := Z_k(Z_k^T Z)$ . Hence, for sufficiently large  $k$  we have  $\tilde{Z}_k \in \mathcal{M}_n$  and  $\tilde{Z}_k \in \mathcal{M}(L_k)$  follows. This proves (ii) and (iii) follows from (ii) and (i). In order to prove the compactness of  $\mathcal{L}_n$ , consider a sequence  $L_k \in \mathcal{L}_n$  together with basis matrices  $Z_k \in \mathcal{M}^{\text{orth}}(L_k)$ . By possibly passing to a subsequence we may assume that  $Z_k$  converges to some  $Z$ . Since  $Z^T Z = \lim_{k \rightarrow \infty} Z_k^T Z_k = I$ , we conclude  $Z \in \mathcal{M}_n$  and  $\text{rge } Z = \lim_{k \rightarrow \infty} L_k \in \mathcal{L}_n$  by (i). Hence the metric space  $\mathcal{L}_n$  is (sequentially) compact. Finally, by [34, Example 5.35] there holds  $L_k \rightarrow L$  in the sense of Painlevé-Kuratowski convergence if and only if the projections  $P_{L_k}$  on  $L_k$  converge graphically to the projection  $P_L$  on  $L$ . Since the projections on subspaces in  $\mathcal{L}_n$  are linear mappings with norm equal to 1, graphical convergence of  $P_{L_k}$  to  $P_L$  is equivalent to uniform convergence  $\lim_{k \rightarrow \infty} \|P_{L_k} - P_L\| = 0$  by [34, Theorems 5.43, 5.44].  $\square$

We treat every element of  $\mathbb{R}^{2n}$  as a column vector. In order to keep our notation simple we write  $(u, v)$  instead of  $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^{2n}$  when this does not lead to confusion. In order to refer to the components of the vector  $z = \begin{pmatrix} u \\ v \end{pmatrix}$  we set  $\pi_1(z) := u$ ,  $\pi_2(z) := v$ .

Let  $L \in \mathcal{L}_n$  and consider  $Z \in \mathcal{M}(L)$ , which can be written in the form  $Z = \begin{pmatrix} A \\ B \end{pmatrix}$ . But we will rather write it as  $Z = (A, B)$ ; thus  $\text{rge } (A, B) := \{(Au, Bu) \mid u \in \mathbb{R}^n\} \doteq \left\{ \begin{pmatrix} Au \\ Bu \end{pmatrix} \mid u \in \mathbb{R}^n \right\} = L$ . Similarly as before, we will also use  $\pi_1(Z) := A$ ,  $\pi_2(Z) := B$  for referring to the two  $n \times n$  parts of  $Z$ .

Further, for every  $L \in \mathcal{L}_n$  we define

$$L^* := \{(-v^*, u^*) \mid (u^*, v^*) \in L^\perp\}, \tag{15}$$

where  $L^\perp$  denotes as usual the orthogonal complement of  $L$ . Note that

$$(L^*)^\perp = \{(v, u) \mid \langle v, -v^* \rangle + \langle u, u^* \rangle = 0 \ \forall (u^*, v^*) \in L^\perp\} = \{(v, u) \mid (u, -v) \in (L^\perp)^\perp\}$$

and therefore

$$(L^*)^* = \{(u, v) \mid (v, -u) \in (L^*)^\perp\} = \{(u, v) \mid (-u, -v) \in L\} = L.$$

We denote by  $S_n$  the  $2n \times 2n$  orthogonal matrix

$$S_n := \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix},$$

so that  $L^* = S_n L^\perp$ . If  $P$  represents the orthogonal projection on  $L$  then  $I - P$  is the orthogonal projection on  $L^\perp$  and  $S_n(I - P)S_n^T$  is the orthogonal projection on  $L^*$ . Given two subspaces  $L_1, L_2 \in \mathcal{L}_n$  with orthogonal projections  $P_1, P_2$ , we obtain

$$d_{\mathcal{L}}(L_1^*, L_2^*) = \|S_n(I - P_1)S_n^T - S_n(I - P_2)S_n^T\| = \|S_n(I - P_1 - (I - P_2))S_n^T\| = \|P_1 - P_2\| = d_{\mathcal{L}}(L_1, L_2).$$

Thus the mapping  $L \mapsto L^*$  defines an isometry on  $\mathcal{L}_n$  and a sequence  $(L_k)$  converges in  $\mathcal{L}_n$  to some  $L$  if and only if the sequence  $(L_k^*)$  converges to  $L^*$ .

Consider the following relation between the graphical derivative and differentiability in case of single-valued mappings.

**Lemma 3.2.** *Consider  $f : U \rightarrow \mathbb{R}^n$  with  $U \subseteq \mathbb{R}^n$  open and a point  $x \in U$ . Then one has:*

- (i) *If  $f$  is Fréchet differentiable at  $x$ , then  $DF(x)$  is a single-valued linear mapping,  $DF(x)(u) = \nabla f(x)u$ ,  $u \in \mathbb{R}^n$ , and consequently  $T_{\text{gph } f}(x, f(x)) = \text{rge}(I, \nabla f(x)) \in \mathcal{L}_n$ .*
- (ii) *Conversely, if  $T_{\text{gph } f}(x, f(x)) \in \mathcal{L}_n$  and  $f$  is calm at  $x$ , i.e., there is some  $\kappa \geq 0$  such that the estimate  $\|f(x') - f(x)\| \leq \kappa \|x' - x\|$  holds for all  $x'$  sufficiently close to  $x$ , then  $f$  is Fréchet differentiable at  $x$ .*

**Proof.** The statement (i) follows immediately from [34, Exercise 9.25]. In order to show (ii), we first prove that there is an  $n \times n$  matrix  $A$  such that  $T_{\text{gph } f}(x, f(x)) = \text{rge}(I, A)$ . Considering any  $Z \in \mathcal{M}(T_{\text{gph } f}(x, f(x)))$ , we will show that  $B := \pi_1(Z)$  is nonsingular. Assuming on the contrary that  $B$  is singular, there is some  $p \neq 0$  with  $Bp = 0$ . Then  $v := \pi_2(Z)p \neq 0$  because otherwise  $Zp = 0$  which is not possible. Hence  $(0, v) = Zp \in T_{\text{gph } f}$  and there exists sequences  $t_k \downarrow 0$  and  $(u_k, v_k) \rightarrow (u, v)$  such that  $f(x) + t_k v_k = f(x + t_k u_k) \forall k$  implying  $t_k \|v_k\| = \|f(x + t_k u_k) - f(x)\| \leq \kappa t_k \|u_k\|$  and  $\|v\| = \lim_{k \rightarrow \infty} \|u_k\| = 0$ , a contradiction. Hence  $B$  is nonsingular and we obtain  $T_{\text{gph } f}(x, f(x)) = \text{rge}(B, \pi_2(Z)) = \text{rge}(I, \pi_2(Z)B^{-1})$  proving our claim with  $A = \pi_2(Z)B^{-1}$ . Hence  $Df(x)u = Au$ ,  $u \in \mathbb{R}^n$  and the assertion follows once more from [34, Exercise 9.25].  $\square$

Note that, when  $f : U \rightarrow \mathbb{R}^n$ ,  $U \subseteq \mathbb{R}^n$  open, is Fréchet differentiable at  $u \in U$ , then we even have  $T_{\text{gph } f}(u, f(u)) = \text{Lim}_{t \downarrow 0} t^{-1}(\text{gph } f - (u, f(u)))$ .

We now introduce new generalized derivatives for set-valued mappings. We confine ourselves to the particular case  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and, as we will see in the sequel, this restriction still permits a considerable number of applications.

**Definition 3.3.** Consider a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ .

1. We say that  $F$  is *graphically smooth of dimension  $n$*  at  $(x, y) \in \text{gph } F$ , if  $T_{\text{gph } F}(x, y) = \text{gph } DF(x, y) \in \mathcal{L}_n$ . Further we denote by  $\mathcal{O}_F$  the set of all points where  $F$  is graphically smooth of dimension  $n$ .
2. Associate with  $F$  the four mappings  $\widehat{\mathcal{S}}F, \widehat{\mathcal{S}}^*F, \mathcal{S}F, \mathcal{S}^*F$ , all of which map  $\text{gph } F \rightrightarrows \mathcal{L}_n$  and are given by

$$\begin{aligned} \widehat{\mathcal{S}}F(x, y) &:= \begin{cases} \{\text{gph } DF(x, y)\} & \text{if } (x, y) \in \mathcal{O}_F, \\ \emptyset & \text{else,} \end{cases} \\ \widehat{\mathcal{S}}^*F(x, y) &:= \begin{cases} \{\text{gph } DF(x, y)^*\} & \text{if } (x, y) \in \mathcal{O}_F, \\ \emptyset & \text{else,} \end{cases} \\ \mathcal{S}F(x, y) &:= \text{Lim sup}_{(u,v) \xrightarrow{\text{gph } F} (x,y)} \widehat{\mathcal{S}}F(u, v) \\ &= \{L \in \mathcal{X}_n \mid \exists(x_k, y_k) \xrightarrow{\mathcal{O}_F}(x, y) : \lim_{k \rightarrow \infty} d_{\mathcal{X}}(L, \text{gph } DF(x_k, y_k)) = 0\}, \\ \mathcal{S}^*F(x, y) &= \text{Lim sup}_{(u,v) \xrightarrow{\text{gph } F} (x,y)} \widehat{\mathcal{S}}^*F(u, v) \\ &= \{L \in \mathcal{X}_n \mid \exists(x_k, y_k) \xrightarrow{\mathcal{O}_F}(x, y) : \lim_{k \rightarrow \infty} d_{\mathcal{X}}(L, \text{gph } DF(x_k, y_k)^*) = 0\}. \end{aligned}$$

- 3. (a) We say that  $F$  has the *SCD (subspace containing derivative) property at*  $(x, y) \in \text{gph } F$ , if  $\mathcal{S}^*F(x, y) \neq \emptyset$ .
- (b) We say that  $F$  has the *SCD property around*  $(x, y) \in \text{gph } F$ , if there is a neighborhood  $W$  of  $(x, y)$  such that  $F$  has the SCD property at every  $(x', y') \in \text{gph } F \cap W$ .
- (c) Finally, we call  $F$  an *SCD mapping* if  $F$  has the SCD property at every point of its graph.

Apart from the collections  $\mathcal{S}F(x, y)$  and  $\mathcal{S}^*F(x, y)$  of subspaces we will sometimes use the unions

$$\bigcup \mathcal{S}F(x, y) := \bigcup_{L \in \mathcal{S}F(x,y)} L, \quad \bigcup \mathcal{S}^*F(x, y) := \bigcup_{L \in \mathcal{S}^*F(x,y)} L. \tag{16}$$

**Remark 3.4.** By definition of the regular coderivative there holds

$$\text{gph } \widehat{D}^*F(x, y) = \text{gph } DF(x, y)^*, \quad (x, y) \in \mathcal{O}_F.$$

**Remark 3.5.** Since  $L \mapsto L^*$  is an isometry on  $\mathcal{X}_n$  and  $(L^*)^* = L$ , we have

$$\mathcal{S}^*F(x, y) = \{L^* \mid L \in \mathcal{S}F(x, y)\}, \quad \mathcal{S}F(x, y) = \{L^* \mid L \in \mathcal{S}^*F(x, y)\}.$$

Hence,  $F$  has the SCD property at  $(x, y) \in \text{gph } F$  if and only if  $\mathcal{S}F(x, y) \neq \emptyset$ .

Since we consider convergence in the compact metric space  $\mathcal{X}_n$ , we obtain readily the following result.

**Lemma 3.6.** *A mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  has the SCD property at  $(x, y) \in \text{gph } F$  if and only if  $(x, y) \in \text{cl } \mathcal{O}_F$ . Further,  $F$  is an SCD mapping if and only if  $\text{cl } \mathcal{O}_F = \text{cl } \text{gph } F$ , i.e.,  $F$  is graphically smooth of dimension  $n$  at the points of a dense subset of its graph.*

The name ‘‘SCD property’’ is motivated by the following statement.

**Lemma 3.7.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(x, y) \in \text{gph } F$ . Then  $\bigcup \mathcal{S}^*F(x, y) \subseteq \text{gph } D^*F(x, y)$ .*

**Proof.** Let  $L \in \mathcal{S}^*F(x, y)$  and consider a sequence  $(x_k, y_k, L_k) \rightarrow (x, y, L)$  with  $(x_k, y_k) \in \mathcal{O}_F$  and  $L_k := (\text{gph } DF(x_k, y_k))^* \in \widehat{\mathcal{S}}^*F(x_k, y_k)$ . By Remark 3.4 we have  $L_k = \text{gph } \widehat{D}^*F(x_k, y_k)$ . Consider  $Z_k \in \mathcal{M}^{\text{orth}}(L_k)$ . By possibly passing to a subsequence the matrices  $Z_k$  converge to some  $Z$  and  $L = \text{rge } Z$  by Lemma 3.2. Taking into account  $Z_k p \in L_k = \text{gph } \widehat{D}^*F(x_k, y_k)$ , we obtain  $Zp = \lim_{k \rightarrow \infty} Z_k p \in$

$\text{gph } D^*F(x, y) \forall p \in \mathbb{R}^n$  by the Definitions 2.1, 2.2 showing that  $L \subseteq \text{gph } D^*F(x, y)$ . Since this holds for every  $L \in \mathcal{S}^*F(x, y)$ , the assertion follows.  $\square$

We will now show that the primal subspaces  $L \in \mathcal{S}F(x, y)$  also belong to the graph of some suitable generalized derivative mapping. Consider the following definition.

**Definition 3.8.**

1. Let  $A \subset \mathbb{R}^n$  and let  $\bar{x} \in A$ . The *outer limiting tangent cone* to  $A$  at  $\bar{x}$  is defined as

$$T_A^\sharp(\bar{x}) := \text{Lim sup}_{x \xrightarrow{A} \bar{x}} T_A(x) = \text{Lim sup}_{x \xrightarrow{A} \bar{x}} \left( \text{Lim sup}_{t \downarrow 0} \frac{A - x}{t} \right) \tag{17}$$

2. Consider a multifunction  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  and let  $(\bar{x}, \bar{y}) \in \text{gph } F$ . The *outer limiting graphical derivative* of  $F$  at  $(\bar{x}, \bar{y})$  is the multifunction  $D^\sharp F(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  given by

$$\text{gph } D^\sharp F(\bar{x}, \bar{y}) = T_{\text{gph } F}^\sharp(\bar{x}, \bar{y}).$$

**Remark 3.9.** Comparing (17) with the definition of the paratingent cone  $T_A^P(\bar{x})$  it follows that  $T_A^\sharp(\bar{x}) \subseteq T_A^P(\bar{x})$  and therefore  $D^\sharp F(\bar{x}, \bar{y})(u) \subseteq D_*F(\bar{x}, \bar{y})(u)$ ,  $u \in \mathbb{R}^n$ .

Using similar arguments as in the proof of Lemma 3.7 one obtains the following result.

**Lemma 3.10.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(x, y) \in \text{gph } F$ . Then  $\bigcup \mathcal{S}F(x, y) \subseteq \text{gph } D^\sharp F(x, y)$ .*

For single-valued mappings the constructions of Definition 3.3 are related to the B-subdifferential.

**Lemma 3.11.** *Let  $U \subset \mathbb{R}^n$  be open and let  $f : U \rightarrow \mathbb{R}^n$  be continuous. Then for every  $x \in U$  there holds*

$$\mathcal{S}f(x) := \mathcal{S}(x, f(x)) \supseteq \{\text{rge}(I, A) \mid A \in \overline{\nabla}f(x)\}, \tag{18}$$

$$\mathcal{S}^*f(x) := \mathcal{S}^*(x, f(x)) \supseteq \{\text{rge}(I, A^T) \mid A \in \overline{\nabla}f(x)\}. \tag{19}$$

*If  $f$  is Lipschitz continuous near  $x$ , these inclusions hold with equality and  $f$  has the SCD property around  $x$ .*

**Proof.** Consider  $x \in U$  and  $A \in \overline{\nabla}f(x)$  together with sequences  $x_k \rightarrow x$  and  $\nabla f(x_k) \rightarrow A$ . Then for each  $k$  we have  $T_{\text{gph } f}(x_k, f(x_k)) = \text{rge}(I, \nabla f(x_k)) =: L_k \in \mathcal{L}_n$  by Lemma 3.2 implying that  $(x_k, f(x_k)) \in \mathcal{O}_f$  and  $\widehat{\mathcal{S}}f(x_k) = \{L_k\}$ . Thus the subspaces  $L_k$  converge in  $\mathcal{L}_n$  to  $\text{rge}(I, A) \in \mathcal{S}f(x)$  by Lemma 3.1(i). This proves (18). By taking into account the identity  $\text{rge}(I, A)^\perp = \text{rge}(-A^T, I)$ , it follows that  $\text{rge}(I, A)^* = \text{rge}(I, A^T)$  verifying (19). Now assume that  $f$  is Lipschitzian near  $x$  and consider  $L \in \mathcal{S}f(x)$  together with a sequence  $(x_k, f(x_k)) \xrightarrow{\mathcal{O}_f} (x, f(x))$  such that  $L_k := T_{\text{gph } f}(x_k, f(x_k)) \xrightarrow{\mathcal{L}_n} L$ . By Lemma 3.2 we conclude that  $f$  is differentiable at  $x_k$  and  $L_k = \text{rge}(I, \nabla f(x_k))$ . By Lipschitz continuity of  $f$  the derivatives  $\nabla f(x_k)$  are bounded. Hence, by possibly passing to a subsequence, we can assume that  $\nabla f(x_k)$  converges to some  $A \in \overline{\nabla}f(x)$  and  $L_k \xrightarrow{\mathcal{L}_n} \text{rge}(I, A)$  follows. This proves equality in (18) and equality in (19) easily follows from the identity  $\text{rge}(I, A)^* = \text{rge}(I, A^T)$ . Since  $\overline{\nabla}f(x) \neq \emptyset$  for Lipschitz continuous mappings, the SCD property at  $x$  is established. This also holds for every point sufficiently close to  $x$  and thus  $f$  has the SCD property even around  $x$ .  $\square$

**Remark 3.12.** In particular, every Lipschitz continuous mapping  $f : U \rightarrow \mathbb{R}^n$  with  $U \subseteq \mathbb{R}^n$  open is an SCD mapping. However, the converse is not true: Consider the function  $f(x) = \sqrt{|x|}$  which is an SCD mapping but not Lipschitz continuous.

**Lemma 3.13.** Consider a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(x, y) \in \text{gph } F$ . Then

$$\mathcal{S}F(x, y) = \text{Lim sup}_{(u,v) \xrightarrow{\text{gph } F} (x,y)} \mathcal{S}F(u, v), \quad \mathcal{S}^*F(x, y) = \text{Lim sup}_{(u,v) \xrightarrow{\text{gph } F} (x,y)} \mathcal{S}^*F(u, v).$$

**Proof.** We prove only the first equation. The inclusion  $\mathcal{S}F(x, y) \subseteq \text{Lim sup}_{(u,v) \xrightarrow{\text{gph } F} (x,y)} \mathcal{S}F(u, v) =: S$  follows easily from the definition of  $\mathcal{S}F(x, y)$  together with  $\widehat{\mathcal{S}}F(u, v) \subseteq \mathcal{S}F(u, v)$ ,  $(u, v) \in \text{gph } F$ . In order to show the reverse inclusion, consider  $L \in S$  together with sequences  $(u_k, v_k) \xrightarrow{\text{gph } F} (x, y)$  and  $L_k \in \mathcal{S}F(u_k, v_k)$  with  $L_k \xrightarrow{\mathcal{Z}_n} L$ . By definition, for every  $k$  we can find  $(u'_k, v'_k) \in \mathcal{O}_F$  and  $L'_k \in \widehat{\mathcal{S}}F(u_k, v_k)$  such that  $\|(u_k, v_k) - (u'_k, v'_k)\| \leq \frac{1}{k}$  and  $d_{\mathcal{Z}}(L_k, L'_k) \leq \frac{1}{k}$ . Thus  $(u'_k, v'_k) \xrightarrow{\mathcal{O}_F} (x, y)$  and  $L'_k \xrightarrow{\mathcal{Z}_n} L$  verifying  $L \in \mathcal{S}F(x, y)$ .  $\square$

We now provide some calculus rules.

**Proposition 3.14.** Given a mapping  $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and a mapping  $\Phi : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ , consider the mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  given by

$$\text{gph } F = \{(x, y) \mid \Phi(x, y) \in \text{gph } G\}.$$

Then for every  $(x, y) \in \text{gph } F$  such that  $\Phi$  is continuously differentiable in some neighborhood of  $(x, y)$  and  $\nabla\Phi(x, y)$  is nonsingular, there holds

$$\mathcal{S}F(x, y) = \nabla\Phi(x, y)^{-1} \mathcal{S}G(\Phi(x, y)) := \{\nabla\Phi(x, y)^{-1}L \mid L \in \mathcal{S}G(\Phi(x, y))\}, \tag{20}$$

$$\mathcal{S}^*F(x, y) = S_n \nabla\Phi(x, y)^T S_n^T \mathcal{S}^*G(\Phi(x, y)). \tag{21}$$

**Proof.** By the classical Inverse Function Theorem, there is some open neighborhood  $W$  of  $(x, y)$  such that  $\Phi$  is a one-to-one mapping from  $W$  to the open neighborhood  $\tilde{W} := \Phi(W)$  of  $\Phi(x, y)$  and  $\nabla\Phi(x', y')$  is nonsingular for every  $(x', y') \in W$ . By [34, Exercise 6.7] we have  $T_{\text{gph } F}(x', y') = \nabla\Phi(x', y')^{-1} T_{\text{gph } G}(\Phi(x', y'))$  for all  $(x', y') \in \text{gph } F \cap W$  and it follows that  $\mathcal{O}_G \cap \tilde{W} = \Phi(\mathcal{O}_F \cap W)$ . Consider  $L \in \mathcal{S}F(x, y)$  together with sequences  $(x_k, y_k) \xrightarrow{\mathcal{O}_F} (x, y)$  and  $L_k \in \widehat{\mathcal{S}}F(x_k, y_k)$  converging to  $L$ . Then for all  $k$  sufficiently large we have  $T_{\text{gph } G}(\Phi(x_k, y_k)) = \nabla\Phi(x_k, y_k) T_{\text{gph } F}(x_k, y_k) = \nabla\Phi(x_k, y_k) L_k \in \widehat{\mathcal{S}}G(\Phi(x_k, y_k))$  showing  $\nabla\Phi(x, y)L = \lim_{k \rightarrow \infty} \nabla\Phi(x_k, y_k)L_k \in \mathcal{S}G(\Phi(x, y))$  by Lemma 3.1(iii). This proves that  $\mathcal{S}F(x, y) \subseteq \nabla\Phi(x, y)^{-1} \mathcal{S}G(\Phi(x, y))$ .

To show the reverse inclusion, consider  $L \in \mathcal{S}G(\Phi(x, y))$  together with sequences  $z_k \xrightarrow{\mathcal{O}_G \cap \tilde{W}} \Phi(x, y)$  and  $L_k \in \widehat{\mathcal{S}}G(z_k)$  with  $L_k \rightarrow L$ . It follows that the sequence  $(x_k, y_k) := \Phi^{-1}(z_k) \cap W$  converges to  $(x, y)$  and  $T_{\text{gph } F}(x_k, y_k) = \nabla\Phi(x_k, y_k)^{-1} L_k \in \widehat{\mathcal{S}}F(x_k, y_k)$  implying  $\nabla\Phi(x, y)^{-1}L \in \mathcal{S}F(x, y)$  by Lemma 3.2(iii). Hence  $\nabla\Phi(x, y)^{-1} \mathcal{S}G(\Phi(x, y)) \subseteq \mathcal{S}F(x, y)$  and equation (20) follows. To prove the equation (21), just use Remark 3.5 together with the fact that for any  $L \in \mathcal{Z}_n$  we have  $(\nabla\Phi(x, y)^{-1}L)^\perp = \nabla\Phi(x, y)^T L^\perp$  implying

$$(\nabla\Phi(x, y)^{-1}L)^* = S_n \nabla\Phi(x, y)^T L^\perp = S_n \nabla\Phi(x, y)^T S_n^T L^*. \quad \square$$

**Proposition 3.15.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  have the SCD property at  $(x, y) \in \text{gph } F$  and let  $h : U \rightarrow \mathbb{R}^n$  be continuously differentiable at  $x \in U$  where  $U \subseteq \mathbb{R}^n$  is open. Then  $F+h$  has the SCD property at  $(x, y+h(x))$  and

$$\mathcal{S}(F+h)(x, y+h(x)) = \left\{ \begin{pmatrix} I & 0 \\ \nabla h(x) & I \end{pmatrix} L \mid L \in \mathcal{S}F(x, y) \right\} \quad (22)$$

$$\mathcal{S}^*(F+h)(x, y+h(x)) = \left\{ \begin{pmatrix} I & 0 \\ \nabla h(x)^T & I \end{pmatrix} L \mid L \in \mathcal{S}^*F(x, y) \right\} \quad (23)$$

**Proof.** We have  $\text{gph}(F+h) = \{(u, v+h(u)) \mid (u, v) \in \text{gph} F\} = \{(x, y) \mid (x, y-h(x)) \in \text{gph} F\}$  and the assertion follows from Proposition 3.14 with  $\Phi(x, y) = (x, y-h(x))$ .  $\square$

Next, let us proceed to the large class of graphically Lipschitzian mappings.

**Definition 3.16** (cf. [34, Definition 9.66]). A mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is *graphically Lipschitzian of dimension  $d$*  at  $(\bar{x}, \bar{y}) \in \text{gph} F$  if there is an open neighborhood  $W$  of  $(\bar{x}, \bar{y})$  and a one-to-one mapping  $\Phi$  from  $W$  onto an open subset of  $\mathbb{R}^{n+m}$  with  $\Phi$  and  $\Phi^{-1}$  continuously differentiable, such that  $\Phi(\text{gph} F \cap W)$  is the graph of a Lipschitz continuous mapping  $f : U \rightarrow \mathbb{R}^{n+m-d}$ , where  $U$  is an open set in  $\mathbb{R}^d$ .

In what follows we will refer to the mapping  $\Phi$  as *transformation mapping*.

**Proposition 3.17.** Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y}) \in \text{gph} F$  with transformation mapping  $\Phi$ . Then  $F$  has the SCD property around  $(\bar{x}, \bar{y})$  and for every  $(x, y) \in \text{gph} F$ , sufficiently close to  $(\bar{x}, \bar{y})$ , one has

$$\mathcal{S}F(x, y) = \nabla\Phi(x, y)^{-1} \mathcal{S}f(u) = \left\{ \text{rge} \left[ \nabla\Phi(x, y)^{-1} \begin{pmatrix} I \\ B \end{pmatrix} \right] \mid B \in \overline{\nabla}f(u) \right\}, \quad (24a)$$

$$\mathcal{S}^*F(x, y) = S_n \nabla\Phi(x, y)^T S_n^T \mathcal{S}^*f(u) = \left\{ \text{rge} \left[ S_n \nabla\Phi(x, y)^T S_n^T \begin{pmatrix} I \\ B^T \end{pmatrix} \right] \mid B \in \overline{\nabla}f(u) \right\}, \quad (24b)$$

where  $f$  is as in Definition 3.16 and  $u := \pi_1(\Phi(x, y))$ .

**Proof.** Follows from Proposition 3.14 together with Lemma 3.11.  $\square$

**Remark 3.18.** Note that for a graphically Lipschitzian mapping  $F$  with transformation mapping  $\Phi$  we have  $\Phi(\mathcal{O}_F \cap W) = \mathcal{O}_f$  by the proof of Proposition 3.14, where  $W$  and  $f$  are as in Definition 3.16. At points  $(u, f(u)) \in \mathcal{O}_f$  the mapping  $f$  is Fréchet differentiable at  $u$  by Lemma 3.2(ii) and therefore  $T_{\text{gph} f}(u, f(u)) = \text{Lim}_{t \downarrow 0} t^{-1}(\text{gph} f - (u, f(u)))$ . Since the graphs of  $F$  and  $f$  coincide locally up to a change of coordinates, we may conclude that  $T_{\text{gph} F}(x, y) = \text{Lim}_{t \downarrow 0} t^{-1}(\text{gph} F - (x, y))$ ,  $(x, y) \in \mathcal{O}_F \cap W$ , i.e.,  $F$  is proto-differentiable at these points, cf. [34, Section 8.H].

**Corollary 3.19.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(\bar{x}, \bar{y}) \in \text{gph} F$  be given. Suppose that there is an open neighborhood  $V$  of  $\bar{x}$  and a continuously differentiable mapping  $h : V \rightarrow \mathbb{R}^n$  such that  $F+h$  is strongly metrically regular around  $(\bar{x}, \bar{y}+h(\bar{x}))$ . Then  $F$  is graphically Lipschitzian of dimension  $n$  with transformation mapping  $\Phi(x, y) = (y+h(x), x)$ . Therefore  $F$  has the SCD property around  $(\bar{x}, \bar{y})$  and for every  $(x, y) \in \text{gph} F$  sufficiently close to  $(\bar{x}, \bar{y})$  one has

$$\begin{aligned} \mathcal{S}F(x, y) &= \{ \text{rge} (B, I - \nabla h(x)B) \mid B \in \overline{\nabla}(F+h)^{-1}(y+h(x)) \} \\ &= \begin{pmatrix} 0 & I \\ I & -\nabla h(\bar{x}) \end{pmatrix} \mathcal{S}(F+h)^{-1}(\bar{y}+h(\bar{x}), \bar{x}), \end{aligned} \quad (25a)$$

$$\begin{aligned} \mathcal{S}^*F(x, y) &= \{ \text{rge} (B^T, I - \nabla h(x)^T B^T) \mid B \in \overline{\nabla}(F+h)^{-1}(y+h(x)) \} \\ &= \begin{pmatrix} 0 & I \\ I & -\nabla h(x)^T \end{pmatrix} \mathcal{S}^*(F+h)^{-1}(y+h(x), x). \end{aligned} \quad (25b)$$

**Proof.** By Theorem 2.6 there are open neighborhoods  $U$  of  $\bar{y} + h(\bar{x})$ ,  $W'$  of  $(\bar{y} + h(\bar{x}), \bar{x})$  and a Lipschitz continuous mapping  $f : U \rightarrow \mathbb{R}^n$  such that  $\text{gph } f = \text{gph } (F+h)^{-1} \cap W'$ . Since  $\text{gph } (F+h)^{-1} = \{(y+h(x), x) \mid (x, y) \in \text{gph } F, x \in V\}$ ,  $F$  is graphically Lipschitzian of dimension  $n$  and (25) follows from Proposition 3.17 by taking into account that

$$\nabla\Phi(x, y) = \begin{pmatrix} \nabla h(x) & I \\ I & 0 \end{pmatrix}, \quad \nabla\Phi(x, y)^{-1} = \begin{pmatrix} 0 & I \\ I & -\nabla h(x) \end{pmatrix}, \quad S_n \nabla\Phi(x, y)^T S_n^T = - \begin{pmatrix} 0 & I \\ I & -\nabla h(x)^T \end{pmatrix}. \quad \square$$

Some examples of graphically Lipschitzian mappings  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  of dimension  $n$  were already given in [33,27]. Next, we extend these examples and give an explicit description of the subspaces contained in  $\mathcal{S}F(x, y)$  and  $\mathcal{S}^*F(x, y)$ , respectively. Recall that a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is said to be monotone if

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0 \quad \text{for all } (x_i, y_i) \in \text{gph } F, \quad i = 1, 2.$$

It is maximally monotone if, in addition, there holds  $\text{gph } F = \text{gph } T$  for every monotone mapping  $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  with  $\text{gph } F \subset \text{gph } T$ . Next we define several types of local monotonicity.

**Definition 3.20.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(x, y) \in \text{gph } F$ . We say the following:

1.  $F$  is *locally monotone* at  $(x, y)$  if there is an open neighborhood  $X \times Y$  of  $(x, y)$  such that

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0 \quad \text{for all } (x_i, y_i) \in \text{gph } F \cap (X \times Y), \quad i = 1, 2. \tag{26}$$

It is *locally maximally monotone* if, in addition, there holds  $\text{gph } F \cap (X \times Y) = \text{gph } T \cap (X \times Y)$  for every monotone mapping  $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  with  $\text{gph } F \cap (X \times Y) \subseteq \text{gph } T$ .

2.  $F$  is *locally (maximally) hypomonotone* at  $(x, y)$  if  $\gamma I + F$  is locally (maximally) monotone at  $(x, \gamma x + y)$  for some  $\gamma \geq 0$ .

Related with maximally monotone operators are the so-called firmly nonexpansive mappings.

**Definition 3.21.**

1. A mapping  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called *firmly nonexpansive* if  $\langle f(u_1) - f(u_2), u_1 - u_2 \rangle \geq \|f(u_1) - f(u_2)\|^2$ ,  $u_1, u_2 \in \mathbb{R}^n$ .
2. An  $n \times n$  matrix  $B$  is called *firmly nonexpansive*, if the linear mapping  $u \rightarrow Bu$  is firmly nonexpansive, i.e.,  $\langle Bv, v \rangle \geq \|Bv\|^2$ ,  $v \in \mathbb{R}^n$ .

Note that an  $n \times n$  matrix  $B$  is firmly nonexpansive if and only if  $\|2B - I\| \leq 1$ , see, e.g., [3, Fact 1.1]. Further, a firmly nonexpansive matrix  $B$  is positive semidefinite and satisfies  $\|B\| \leq 1$  and, when  $B$  is symmetric, these conditions are also sufficient for  $B$  being firmly nonexpansive.

**Proposition 3.22.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be locally maximally monotone at  $(\bar{x}, \bar{y})$ . Then  $F$  is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y})$  with transformation mapping  $\Phi(x, y) = (x + y, x)$  and consequently  $F$  has the SCD property around  $(\bar{x}, \bar{y})$ . Further, for every  $(x, y) \in \text{gph } F$  sufficiently close to  $(\bar{x}, \bar{y})$  and for every subspace  $L \in \mathcal{S}F(x, y)$  there is a firmly nonexpansive  $n \times n$  matrix  $B$  such that  $L = \text{rge } (B, I - B)$  and  $L^* = \text{rge } (B^T, I - B^T)$ .

**Proof.** Let  $F_{1\text{oc}} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be given by  $\text{gph } F_{1\text{oc}} = \text{gph } F \cap (X \times Y)$ , where  $(X \times Y)$  is as in Definition 3.20. Then  $F_{1\text{oc}}$  is monotone, has a maximally monotone extension  $\tilde{F}$ , cf. [34, Proposition 12.6] and  $\text{gph } F \cap (X \times$

$Y) = \text{gph } \tilde{F} \cap (X \times Y)$ . By [2, Corollary 23.8],  $\text{rge } (I + \tilde{F}) = \mathbb{R}^n$  and the resolvent  $f := (I + \tilde{F})^{-1}$  is a single-valued, firmly nonexpansive mapping on  $\mathbb{R}^n$ . Since  $f(\bar{x} + \bar{y}) = \bar{x}$ , we can find an open neighborhood  $U$  of  $\bar{x} + \bar{y}$  such that  $f(u) \in X$  and  $u - f(u) \in Y$  for all  $u \in U$ . It follows that  $(I + F)^{-1}(u) = (I + \tilde{F})^{-1}(u)$  and therefore  $I + F$  is strongly metrically regular around  $(\bar{x}, \bar{x} + \bar{y})$ . Thus  $F$  is graphically Lipschitzian at  $(\bar{x}, \bar{y})$  and has the SCD property around  $(\bar{x}, \bar{y})$  by Corollary 3.19. Now consider a subspace  $L \in \mathcal{S}F(x, y)$ , where  $(x, y)$  is close to  $(\bar{x}, \bar{y})$ . By (25a) there is a matrix  $B \in \overline{\nabla}f(x + y)$  such that  $L = \text{rge } (B, I - B)$ . Since  $f$  is firmly nonexpansive, for every  $(u', f(u')) \in \mathcal{O}_f$  we have  $\langle \nabla f(u')v, v \rangle \geq \|\nabla f(u')v\|^2$  and  $\|Bv\|^2 \leq \langle Bv, v \rangle$ ,  $v \in \mathbb{R}^n$  follows. This completes the proof.  $\square$

If  $F$  is only locally maximally hypomonotone at  $(\bar{x}, \bar{y})$ , it follows that  $(1 + \gamma)I + F$  is strongly metrically regular around  $(\bar{x}, (1 + \gamma)\bar{x} + \bar{y})$  for some  $\gamma \geq 0$ . Hence we obtain the following corollary.

**Corollary 3.23.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be locally maximally hypomonotone at  $(\bar{x}, \bar{y})$ . Then there is some  $\gamma \geq 0$  such that  $F$  is graphically Lipschitzian at  $(\bar{x}, \bar{y})$  of dimension  $n$  with transformation mapping  $\Phi(x, y) = ((1 + \gamma)x + y, x)$  and therefore  $F$  has the SCD property around  $(\bar{x}, \bar{y})$ . For every  $(x, y) \in \text{gph } F$  sufficiently close to  $(\bar{x}, \bar{y})$  and every subspace  $L \in \mathcal{S}F(x, y)$  there is a firmly nonexpansive  $n \times n$  matrix  $B$  such that  $L = \text{rge } (B, I - (1 + \gamma)B)$  and  $L^* = \text{rge } (B^T, I - (1 + \gamma)B^T)$ .*

**Corollary 3.24.** *A mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  which is locally maximally hypomonotone on a dense subset of its graph is an SCD mapping.*

We now consider the subdifferential mapping  $\partial q$  of some lsc function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ .

**Definition 3.25.**

1. A function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is *prox-regular* at  $\bar{x} \in \text{dom } q$  for  $\bar{x}^* \in \partial q(\bar{x})$  if  $q$  is locally lsc around  $\bar{x}$  and there exist  $\varepsilon > 0$  and  $\rho \geq 0$  such that for all  $x', x \in \mathcal{B}_\varepsilon(\bar{x})$  with  $|q(x) - q(\bar{x})| \leq \varepsilon$  one has

$$q(x') \geq q(x) + \langle x^*, x' - x \rangle - \frac{\rho}{2} \|x' - x\|^2 \quad \text{whenever } x^* \in \partial q(x) \cap \mathcal{B}_\varepsilon(\bar{x}^*).$$

When this holds for all  $\bar{x}^* \in \partial q(\bar{x})$ ,  $q$  is said to be prox-regular at  $\bar{x}$ .

2. A function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is called *subdifferentially continuous* at  $\bar{x} \in \text{dom } q$  for  $\bar{x}^* \in \partial q(\bar{x})$  if for any sequence  $(x_k, x_k^*) \xrightarrow{\text{gph } \partial q} (\bar{x}, \bar{x}^*)$  we have  $\lim_{k \rightarrow \infty} q(x_k) = q(\bar{x})$ . When this holds for all  $\bar{x}^* \in \partial q(\bar{x})$ ,  $q$  is said to be subdifferentially continuous at  $\bar{x}$ .

**Proposition 3.26.** *Suppose that  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is prox-regular and subdifferentially continuous at  $\bar{x}$  for  $\bar{x}^* \in \partial q(\bar{x})$ . Then  $\partial q$  is locally maximally hypomonotone at  $(\bar{x}, \bar{x}^*)$  and there is some  $\lambda > 0$  such that  $\partial q$  is graphically Lipschitzian at  $(\bar{x}, \bar{y})$  with transformation mapping  $\Phi(x, x^*) = (x + \lambda x^*, x)$ . Thus  $\partial q$  has the SCD property around  $(\bar{x}, \bar{x}^*)$ . Further, for every  $(x, x^*) \in \text{gph } \partial q$  sufficiently close to  $(\bar{x}, \bar{x}^*)$  one has  $\mathcal{S}^* \partial q(x, x^*) = \mathcal{S} \partial q(x, x^*)$  and for every  $L \in \mathcal{S} \partial q(x, x^*)$  there is a symmetric positive semidefinite  $n \times n$  matrix  $B$  such that  $L = L^* = \text{rge } (B, \frac{1}{\lambda}(I - B))$ .*

**Proof.** Let  $\tilde{q}(x) := q(x) - \langle \bar{x}^*, x \rangle$ . Then  $\partial \tilde{q}(\cdot) = \partial q(\cdot) - \bar{x}^*$  and  $\tilde{q}$  is prox-regular at  $\bar{x}$  for 0. By the definition of prox-regularity we have

$$\tilde{q}(x') \geq \tilde{q}(\bar{x}) - \frac{\rho}{2} \|x' - \bar{x}\|^2 \quad \forall x' \in \mathcal{B}_\varepsilon(\bar{x})$$

for some  $\varepsilon > 0$  and some  $\rho \geq 0$ . Hence the function  $\hat{q} := \tilde{q} + \delta_{\mathcal{B}_\varepsilon}$  fulfills the baseline assumption of [27, Section 4]. By [27, Proposition 4.8] the subdifferential mapping  $\partial \hat{q}$  is locally maximally hypomonotone around  $(\bar{x}, 0)$



and, by the proof of [27, Theorem 4.7], for any  $\lambda \in (0, \frac{1}{\rho})$  the mapping  $\partial\hat{q}$  is graphically Lipschitzian with transformation mapping  $\Phi(x, x^*) = (x + \lambda x^*, x)$ . Moreover, by [27, Theorem 4.4],  $(I + \lambda\partial\hat{q})^{-1}$  is locally monotone at  $(\bar{x}, \bar{x})$  and there holds

$$\nabla e_\lambda \hat{q}(u) = \frac{1}{\lambda}(I - (I + \lambda\partial\hat{q})^{-1})(u) \tag{27}$$

for all  $u$  sufficiently close to  $\bar{x}$ , where

$$e_\lambda \hat{q}(y) := \inf_x \left\{ \frac{1}{2\lambda} \|x - y\|^2 + \hat{q}(x) \right\}$$

denotes the Moreau envelope of  $\hat{q}$ . Consider a pair  $(x, \hat{x}^*)$  close to  $(\bar{x}, 0)$  and a subspace  $L \in \mathcal{S}\partial\hat{q}(x, \hat{x}^*)$ . According to Corollary 3.23, there is a matrix  $B \in \overline{\nabla}(I + \lambda\partial\hat{q})^{-1}(x + \lambda\hat{x}^*)$  with  $L = \nabla\Phi(x, \hat{x}^*)^{-1} \text{rge}(I, B) = \text{rge}(B, \frac{1}{\lambda}(I - B))$ , where we have taken into account

$$\nabla\Phi(x, \hat{x}^*) = \begin{pmatrix} I & \lambda I \\ I & 0 \end{pmatrix}, \quad \nabla\Phi(x, \hat{x}^*)^{-1} = \begin{pmatrix} 0 & I \\ \frac{1}{\lambda}I & -\frac{1}{\lambda}I \end{pmatrix}.$$

Since  $(I + \lambda\partial\hat{q})^{-1}$  is locally monotone at  $(\bar{x}, \bar{x})$ , it follows that  $B$  is positive semidefinite. Further, by (27) we have  $\frac{1}{\lambda}(I - B) \in \overline{\nabla}(\nabla e_\lambda \hat{q})(x + \lambda\hat{x}^*)$ . By [34, Theorem 13.52],  $\overline{\nabla}(\nabla e_\lambda \hat{q})(x + \lambda\hat{x}^*)$  consists of symmetric matrices and consequently  $B$  is symmetric. Since  $L^\perp = \text{rge}(\frac{1}{\lambda}(I - B), -B)$ , we obtain  $L^* = L$  and  $\mathcal{S}\partial\hat{q}(x, \hat{x}^*) = \mathcal{S}\partial\hat{q}(x, \hat{x}^*)$  follows. Now, by taking into account that  $\partial\tilde{q}$  and  $\partial\hat{q}$  coincide near  $\bar{x}$  and  $\partial q$  differs from  $\partial\tilde{q}$  only by the constant  $\bar{x}^*$ , it follows that all the shown properties do not hold only for  $\partial\hat{q}$  but also for  $\partial q$ .  $\square$

**Corollary 3.27.** *For every lsc function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  which is prox-regular and subdifferentially continuous at  $x$  for  $x^*$  on a dense subset of  $\text{gph } \partial q$ , its subdifferential mapping  $\partial q$  is an SCD mapping.*

Clearly, every lsc convex function is prox-regular and subdifferentially continuous at all points of its domain. Further, the proof of Proposition 3.26 holds true with  $\lambda = 1$  and  $(I + \partial q)^{-1}$  is firmly nonexpansive and therefore  $\|B\| \leq 1 \forall B \in \overline{\nabla}(I + \partial\hat{q})^{-1}(\bar{x})$ . Thus we obtain the following corollary.

**Corollary 3.28.** *For every lsc proper convex function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  the subdifferential mapping  $\partial q$  is graphically Lipschitzian of dimension  $n$  at every point  $(x, x^*)$  of its graph. Hence  $\partial q$  is an SCD mapping and for every  $(x, x^*) \in \text{gph } \partial q$  and every  $L \in \mathcal{S}^*\partial q(x, x^*) = \mathcal{S}\partial q(x, x^*)$  there is a symmetric positive semidefinite  $n \times n$  matrix  $B$  with  $\|B\| \leq 1$  such that  $L = \text{rge}(B, I - B)$ .*

**Example 3.29.** Let  $C \subset \mathbb{R}^n$  be a convex polyhedral set and consider  $q = \delta_C$  so that  $\partial q = N_C$ . By the well-known reduction Lemma, see, e.g., [10, Lemma 2E.4], we have  $T_{\text{gph } N_C}(x, x^*) = \text{gph } N_{\mathcal{K}_C(x, x^*)}$ , where  $\mathcal{K}_C(x, x^*) := T_C(x) \cap [x^*]^\perp$  denotes the critical cone to  $C$  at  $x$  for  $x^*$ . Thus  $N_C$  is graphically smooth of dimension  $n$  at  $(x, x^*)$  if and only if  $\mathcal{K}_C(x, x^*)$  is a subspace and in this case we have  $\widehat{\mathcal{S}}N_C(x, x^*) = K_C(x, x^*) \times K_C(x, x^*)^\perp = \widehat{\mathcal{S}}^*N_C(x, x^*)$ . Given  $(\bar{x}, \bar{x}^*) \in \text{gph } N_C$ , by [10, Lemma 4H.2], for every sufficiently small neighborhood  $W$  of  $(\bar{x}, \bar{x}^*)$ , the collection of all critical cones  $\mathcal{K}_C(x, x^*)$ ,  $(x, x^*) \in W$  coincides with the collection of all sets of the form  $\mathcal{F}_1 - \mathcal{F}_2$ , where  $\mathcal{F}_1, \mathcal{F}_2$  are faces of  $\mathcal{K}_C(\bar{x}, \bar{x}^*)$  with  $\mathcal{F}_2 \subseteq \mathcal{F}_1$ . Since  $\mathcal{F}_1 - \mathcal{F}_2$  is a subspace if and only if  $\mathcal{F}_1 = \mathcal{F}_2$  and  $\mathcal{K}_C(\bar{x}, \bar{x}^*)$  has only finitely many faces, we obtain

$$\mathcal{S}N_C(\bar{x}, \bar{x}^*) = \mathcal{S}^*N_C(\bar{x}, \bar{x}^*) = \{(\mathcal{F} - \mathcal{F}) \times (\mathcal{F} - \mathcal{F})^\perp \mid \mathcal{F} \text{ is face of } \mathcal{K}_C(\bar{x}, \bar{x}^*)\}. \tag{28}$$

Of course, for every face  $\mathcal{F}$  of  $\mathcal{K}_C(\bar{x}, \bar{x}^*)$  we have

$$(\mathcal{F} - \mathcal{F}) \times (\mathcal{F} - \mathcal{F})^\perp = \text{rge}(B, I - B),$$

where  $B$  represents the orthogonal projection on  $\mathcal{F} - \mathcal{F}$ .

Let us compare the representation (28) with the limiting coderivative  $D^*N_C(\bar{x}, \bar{x}^*)$ . It was shown in [9, Proof of Theorem 2] that  $N_{\text{gph}N_C}(\bar{x}, \bar{x}^*)$  is the union of all product sets  $K^\circ \times K$  associated with cones  $K$  of the form  $\mathcal{F}_1 - \mathcal{F}_2$ , where  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are closed faces of the critical cone  $K_C(\bar{x}, \bar{x}^*)$  satisfying  $\mathcal{F}_2 \subset \mathcal{F}_1$ . Thus,  $\text{gph}D^*N_C(\bar{x}, \bar{x}^*)$  is the union of all respective sets of the form  $(\mathcal{F}_2 - \mathcal{F}_1) \times (\mathcal{F}_1 - \mathcal{F}_2)^\circ$  and we see that  $\mathcal{S}^*N_C(\bar{x}, \bar{x}^*)$  has a simpler structure than the limiting coderivative  $D^*N_C(\bar{x}, \bar{x}^*)$  whenever the critical cone  $K_C(\bar{x}, \bar{x}^*)$  is not a subspace.

Given a function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ ,  $\mathcal{S}^*\partial q(x, x^*)$  amounts to a generalized derivative of the subgradient mapping and constitutes therefore some generalized second-order derivative of  $q$ . In the framework of a study of sufficient conditions for local optimality, Rockafellar [35] has introduced another type of generalized second-order derivative as an epigraphical limit of certain second-order subderivatives.

**Definition 3.30** ([35]).

1. A function  $\phi : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is called a *generalized quadratic form*, if  $\phi(0) = 0$  and the subgradient mapping  $\partial\phi$  is *generalized linear*, i.e.,  $\text{gph}\partial\phi$  is a subspace of  $\mathbb{R}^n \times \mathbb{R}^n$ .
2. A function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is called *generalized twice differentiable* at  $x$  for a subgradient  $x^* \in \partial q(x)$ , if it is twice epi-differentiable at  $x$  for  $x^*$  with the second-order subderivative  $d^2q(x, x^*)$  being a generalized quadratic form.
3. Given a function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  and a pair  $(x, x^*) \in \text{gph}\partial q$ , the *quadratic bundle* of  $q$  at  $x$  for  $x^*$  is defined by

$$\text{quad}q(x, x^*) := \left[ \begin{array}{l} \text{the collection of generalized quadratic forms } \phi \text{ for which} \\ \exists(x_k, x_k^*) \rightarrow (x, x^*) \text{ with } q \text{ generalized twice differentiable at} \\ x_k \text{ for } x_k^* \text{ and such that the generalized quadratic forms } \phi_k = \\ d^2q(x_k, x_k^*) \text{ converge epigraphically to } \phi. \end{array} \right.$$

We establish now a strong relationship between  $\mathcal{S}^*\partial q$  and  $\text{quad}q$  for prox-regular and subdifferentially continuous functions  $q$ . We start with the following lemma.

**Lemma 3.31.** *Let  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  be prox-regular and subdifferentially continuous at  $x$  for  $x^* \in \partial q(x)$ . Then  $(x, x^*) \in \mathcal{O}_{\partial q}$  if and only if  $q$  is generalized twice differentiable at  $x$  for  $x^*$ , and in this case one has*

$$\widehat{\mathcal{S}^*\partial}q(x, x^*) = \{\text{gph}\partial(\frac{1}{2}d^2q(x, x^*))\}. \tag{29}$$

**Proof.** By [34, Theorem 13.40], the subgradient mapping  $\partial q$  is proto-differentiable at  $(x, x^*)$  if and only if  $q$  is twice epi-differentiable at  $x$  for  $x^*$ , and then

$$D(\partial q)(x, x^*) = \partial(\frac{1}{2}d^2q(x, x^*)).$$

By Proposition 3.26,  $\partial q$  is graphically Lipschitzian of order  $n$  at  $(x, x^*)$ . Hence, if  $(x, x^*) \in \mathcal{O}_{\partial q}$  then  $\partial q$  is proto-differentiable at  $(x, x^*)$  by Remark 3.18 and consequently  $q$  is twice epi-differentiable at  $x$  for  $x^*$  and  $\text{gph}\partial(\frac{1}{2}d^2q(x, x^*)) = \text{gph}D(\partial q)(x, x^*) = T_{\text{gph}\partial q}(x, x^*)$  is a subspace. This verifies that  $q$  is generalized twice differentiable at  $x$  for  $x^*$ . Conversely, if  $q$  is generalized twice differentiable at  $x$  for  $x^*$ ,

then  $T_{\text{gph } \partial q}(x, x^*) = \text{gph } \partial(\frac{1}{2}d^2q(x, x^*))$  is a subspace. The dimension of this subspace must be  $n$  because  $\partial q$  is graphically Lipschitzian of dimension  $n$ , and  $(x, x^*) \in \mathcal{O}_{\partial q}$  follows.  $\square$

We will also make use of the following variant of Attouch’s theorem, see, e.g., [34, Theorem 12.35], which states the connection between graphical convergence of subdifferential mappings and epi-convergence of the functions themselves, when the functions are convex.

**Lemma 3.32.** *Let  $\phi_k : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}, k \in \mathbb{N}$  and  $\phi : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  be proper lsc functions and assume that there is some  $\rho \geq 0$  such that the functions  $\hat{\phi}_k := \phi_k + \rho\|\cdot\|^2$  are convex. Then the following two statements are equivalent.*

- (i) *The functions  $\phi_k$  epi-converge to  $\phi$ .*
- (ii) *The mappings  $\partial\phi_k$  converge graphically to  $\partial\phi$ ,  $\hat{\phi} := \phi + \rho\|\cdot\|^2$  is convex and there is some sequence  $(x_k, x_k^*)$  converging to some  $(\bar{x}, \bar{x}^*) \in \text{gph } \partial\phi$  such that  $(x_k, x_k^*) \in \text{gph } \partial\phi_k \forall k$  and  $\lim_{k \rightarrow \infty} \phi_k(x_k) = \phi(\bar{x})$ .*

**Proof.** Observe that assertion (i) is equivalent to epi-convergence of  $\hat{\phi}_k \rightarrow \hat{\phi}$  by [34, Exercise 7.8], and in this case the epigraphical limit  $\hat{\phi}$  is convex, cf. [34, Theorem 7.17]. Similarly,  $\partial\phi_k$  converges graphically to  $\partial\phi$  if and only if  $\partial\hat{\phi}_k$  converges graphically to  $\partial\hat{\phi}$ . Indeed, since  $\partial\hat{\phi}_k(x) = \partial\phi_k(x) + 2\rho x$ , we easily obtain

$$\begin{aligned} \text{Lim sup}_{k \rightarrow \infty} \text{gph } \hat{\phi}_k &= \{(x, x^* + 2\rho x) \mid (x, x^*) \in \text{Lim sup}_{k \rightarrow \infty} \text{gph } \phi_k\}, \\ \text{Lim inf}_{k \rightarrow \infty} \text{gph } \hat{\phi}_k &= \{(x, x^* + 2\rho x) \mid (x, x^*) \in \text{Lim inf}_{k \rightarrow \infty} \text{gph } \phi_k\} \end{aligned}$$

and the claim follows. Finally we have  $\phi(\bar{x}) = \lim_{k \rightarrow \infty} \phi_k(x_k)$  for some sequence  $(x_k, x_k^*) \rightarrow (\bar{x}, \bar{x}^*) \in \text{gph } \partial\phi$  with  $(x_k, x_k^*) \in \text{gph } \partial\phi_k$ , if and only if  $\hat{\phi}(\bar{x}) = \lim_{k \rightarrow \infty} \hat{\phi}_k(x_k)$  and  $\text{gph } \partial\hat{\phi}_k \ni (x_k, x_k^* + 2\rho x_k) \rightarrow (\bar{x}, \bar{x}^* + 2\rho\bar{x}) \in \text{gph } \partial\hat{\phi}$ . Now the equivalence between (i) and (ii) follows from Attouch’s theorem [34, Theorem 12.35] applied to the convex functions  $\hat{\phi}_k$  and  $\hat{\phi}$ .  $\square$

**Proposition 3.33.** *Suppose that  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is prox-regular at  $\bar{x}$  for  $\bar{x}^* \in \partial q(\bar{x})$ . Further assume that for all  $(x, x^*) \in \text{gph } \partial q$  sufficiently close to  $(\bar{x}, \bar{x}^*)$  the function  $q$  is subdifferentially continuous at  $x$  for  $x^*$ . Then*

$$\mathcal{S}^* \partial q(\bar{x}, \bar{x}^*) = \mathcal{S} \partial q(\bar{x}, \bar{x}^*) = \{\text{gph } \partial(\frac{1}{2}\phi) \mid \phi \in \text{quad } q(\bar{x}, \bar{x}^*)\}.$$

**Proof.** Let  $\varepsilon$  and  $\rho \geq 0$  be as in Definition 3.25. Then we can find an open neighborhood  $W$  of  $(\bar{x}, \bar{x}^*)$  such that  $W \subset \mathcal{B}_{\frac{\varepsilon}{2}}(\bar{x}) \times \mathcal{B}_{\frac{\varepsilon}{2}}(\bar{x}^*)$  and for all  $(x, x^*) \in \text{gph } \partial q \cap W$  the function  $q$  is subdifferentially continuous at  $x$  for  $x^*$  and  $|q(x) - q(\bar{x})| < \frac{\varepsilon}{2}$ . Consider  $(\hat{x}, \hat{x}^*) \in \text{gph } \partial q \cap W$ . Then for all  $x, x' \in \mathcal{B}_{\frac{\varepsilon}{2}}(\hat{x})$  with  $|q(x) - q(\hat{x})| \leq \frac{\varepsilon}{2}$  and all  $x^* \in \partial q(x) \cap \mathcal{B}_{\frac{\varepsilon}{2}}(\hat{x}^*)$  we have  $x, x' \in \mathcal{B}_{\varepsilon}(\bar{x}), |q(x) - q(\bar{x})| \leq \varepsilon$  and  $x^* \in \partial q(x) \cap \mathcal{B}_{\varepsilon}(\bar{x}^*)$  implying  $q(x') \geq q(x) + \langle x^*, x' - x \rangle - \frac{\rho}{2}\|x' - x\|^2$ . Thus  $q$  is prox-regular at  $\hat{x}$  for  $\hat{x}^*$  and therefore graphically Lipschitzian of dimension  $n$  at  $(\hat{x}, \hat{x}^*)$ . Further, by [34, Proposition 13.49] and its proof, we may conclude that  $d^2q(\hat{x}, \hat{x}^*) + \rho\|\cdot\|^2$  is a lsc convex function.

Now consider  $\phi \in \text{quad } q(\bar{x}, \bar{x}^*)$  together with some sequence  $(x_k, x_k^*) \rightarrow (\bar{x}, \bar{x}^*)$  such that  $q$  is generalized twice differentiable at  $x_k$  for  $x_k^*$  and the generalized quadratic forms  $\phi_k := d^2q(x_k, x_k^*)$  converge epigraphically to  $\phi$ . We may assume that  $(x_k, x_k^*) \in \text{gph } \partial q \cap W \forall k$  and thus  $(x_k, x_k^*) \in \mathcal{O}_{\partial q} \cap W$  by Lemma 3.31. By Lemma 3.32,  $\partial\phi_k$  converges graphically to  $\partial\phi$ , i.e.,  $\text{gph } \partial\phi_k \rightarrow \text{gph } \partial\phi$  in the sense of Painlevé-Kuratowski convergence. Now it follows from Lemma 3.1(v) and (29) that  $\text{gph } \partial(\frac{1}{2}\phi) \in \mathcal{S} \partial q(\bar{x}, \bar{x}^*)$ .

Conversely, consider a subspace  $L \in \mathcal{S} \partial q(\bar{x}, \bar{x}^*)$  together with sequences  $(x_k, x_k^*) \in \mathcal{O}_{\partial q} \cap W$  and  $L_k \in \widehat{\mathcal{F}} \partial q(x_k, x_k^*)$  with  $\lim_{k \rightarrow \infty} d_{\mathcal{L}}(L_k, L) = 0$ . According to Proposition 3.26 we can find symmetric positive

semidefinite matrices  $B_k, B$  such that  $L_k = \text{rge}(B_k, \frac{1}{\lambda}(I - B_k))$ ,  $L = \text{rge}(B, \frac{1}{\lambda}(I - B))$  with  $\lambda = 1/(\rho + 1)$ . Now let  $\mathcal{U} := \text{rge } B$  and set  $Q := \frac{1}{\lambda}(B^\dagger - BB^\dagger)$ , where  $B^\dagger$  denotes the Moore-Penrose inverse of  $B$ . Since  $B$  is symmetric and positive semidefinite, so is  $B^\dagger$  as well. Further,  $BB^\dagger B = B$  and  $BB^\dagger = B^\dagger B$  is the orthogonal projection onto  $\text{rge } B$ , so that  $I - B^\dagger B$  is the orthogonal projection onto  $\mathcal{U}^\perp = \ker B$ . Now consider the generalized quadratic form  $\phi(x) := \langle Qx, x \rangle + \delta_{\mathcal{U}}(x)$ . Since  $Q$  is symmetric, we obtain

$$\partial(\frac{1}{2}\phi)(x) = \begin{cases} Qx + \mathcal{U}^\perp = \frac{1}{\lambda}(B^\dagger - BB^\dagger)x + \mathcal{U}^\perp & \text{if } x \in \mathcal{U}, \\ \emptyset & \text{else.} \end{cases}$$

Thus

$$\begin{aligned} \text{gph } \partial(\frac{1}{2}\phi) &= \left\{ (Bp, \frac{1}{\lambda}(B^\dagger - BB^\dagger)Bp + (I - B^\dagger B)v) \mid p, v \in \mathbb{R}^n \right\} \\ &= \left\{ (B(B^\dagger Bp + (I - B^\dagger B)v), \frac{1}{\lambda}(I - B)(B^\dagger Bp + (I - B^\dagger B)v)) \mid p, v \in \mathbb{R}^n \right\} \\ &= \text{rge}(B, \frac{1}{\lambda}(I - B)) = L. \end{aligned}$$

The matrix  $Q + \frac{1}{\lambda}I = \frac{1}{\lambda}B^\dagger + \frac{1}{\lambda}(I - BB^\dagger)$  is positive semidefinite as the sum of two positive semidefinite matrices and therefore the function  $\hat{\phi} := \phi + \frac{1}{\lambda}\|\cdot\|^2 = \phi + (\rho + 1)\|\cdot\|^2$  is convex. For each  $k$ , the function  $\phi_k := d^2q(x_k, x_k^*)$  fulfills  $\text{gph } \partial(\frac{1}{2}q) = L_k$  by Lemma 3.31 and  $\hat{\phi}_k := \phi_k + (\rho + 1)\|\cdot\|^2$  is convex. Since  $(0, 0) \in L_k = \text{gph } \partial(\frac{1}{2}\phi_k)$ , we have  $0 \in \partial\phi_k(0)$ . Further,  $\phi_k(0) = \phi(0) = 0$  and  $0 \in \partial\phi(0)$ . Since convergence of  $L_k$  to  $L$  implies that  $\partial\phi_k$  converges graphically to  $\partial\phi$ , it follows from Lemma 3.32 that  $\phi_k$  converges epigraphically to  $\phi$  and we conclude  $\phi \in \text{quad } q(\bar{x}, \bar{x}^*)$ . Thus  $L \in \{\text{gph } \partial(\frac{1}{2}\phi) \mid \phi \in \text{quad } q(\bar{x}, \bar{x}^*)\}$  verifying  $\mathcal{S}\partial q(\bar{x}, \bar{x}^*) = \{\text{gph } \partial(\frac{1}{2}\phi) \mid \phi \in \text{quad } q(\bar{x}, \bar{x}^*)\}$ . By Proposition 3.26 we have  $\mathcal{S}^*\partial q(\bar{x}, \bar{x}^*) = \mathcal{S}\partial q(\bar{x}, \bar{x}^*)$  and the proof is complete.  $\square$

**Corollary 3.34.** *For every lsc proper convex function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  and for every pair  $(x, x^*) \in \text{gph } \partial q$  we have*

$$\mathcal{S}^*\partial q(x, x^*) = \mathcal{S}\partial q(x, x^*) = \{\text{gph } \partial(\frac{1}{2}\phi) \mid \phi \in \text{quad } q(x, x^*)\}.$$

#### 4. SCD regularity

In this section we present the definition and basic properties of a certain property called SCD regularity, which has various applications as we will demonstrate in the subsequent sections.

**Definition 4.1.**

1. We denote by  $\mathcal{L}_n^{\text{reg}}$  the collection of all subspaces  $L \in \mathcal{L}_n$  such that

$$(y^*, 0) \in L \Rightarrow y^* = 0. \quad (30)$$

2. A mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is called *SCD regular around*  $(x, y) \in \text{gph } F$ , if  $F$  has the SCD property around  $(x, y)$  and

$$(y^*, 0) \in \bigcup \mathcal{S}^*F(x, y) \Rightarrow y^* = 0, \quad (31)$$

i.e.,  $L \in \mathcal{L}_n^{\text{reg}}$  for all  $L \in \mathcal{S}^*F(x, y)$ . Further, we will denote by

$$\text{scd reg } F(x, y) := \sup\{\|y^*\| \mid (y^*, x^*) \in \bigcup \mathcal{S}^* F(x, y), \|x^*\| \leq 1\}$$

the modulus of SCD regularity of  $F$  around  $(x, y)$ .

In the following proposition we state some basic properties of subspaces  $L \in \mathcal{Z}_n^{\text{reg}}$ .

**Proposition 4.2.** *Given a  $2n \times n$  matrix  $Z$ , there holds  $\text{rge } Z \in \mathcal{Z}_n^{\text{reg}}$  if and only if the  $n \times n$  matrix  $\pi_2(Z)$  is nonsingular. Thus, for every  $L \in \mathcal{Z}_n^{\text{reg}}$  there is a unique  $n \times n$  matrix  $C_L$  such that  $L = \text{rge}(C_L, I)$ . Further,  $L^* = \text{rge}(C_L^T, I) \in \mathcal{Z}_n^{\text{reg}}$ ,*

$$\langle x^*, C_L^T v \rangle = \langle y^*, v \rangle \quad \forall (y^*, x^*) \in L \quad \forall v \in \mathbb{R}^n, \tag{32}$$

and

$$\|y^*\| \leq \|C_L\| \|x^*\| \quad \forall (y^*, x^*) \in L. \tag{33}$$

**Proof.** Clearly, if  $\pi_2(Z)$  is nonsingular then  $\text{rge } Z \in \mathcal{Z}_n$ . Further, given  $(y^*, 0) \in \text{rge } Z$ , there is some  $p$  with  $y^* = \pi_1(Z)p$ ,  $0 = \pi_2(Z)p$  implying  $p = y^* = 0$  and therefore  $\text{rge } Z \in \mathcal{Z}_n^{\text{reg}}$ . Conversely, consider  $L \in \mathcal{Z}_n^{\text{reg}}$  and  $Z \in \mathcal{M}(L)$ . Because  $L \in \mathcal{Z}_n$ , the matrix  $Z$  has full column rank  $n$  and therefore there cannot exist  $p \neq 0$  with  $Zp = 0$ . Thus, if  $A := \pi_2(Z)$  were singular, there is some  $0 \neq p \in \mathbb{R}^n$  with  $Ap = 0$  and  $Bp \neq 0$  with  $B := \pi_1(Z)$ , implying  $(Bp, 0) \in L$  and  $Bp \neq 0$  which is not possible because of  $L \in \mathcal{Z}_n^{\text{reg}}$ . This proves that  $A$  is nonsingular and  $L = \text{rge } ZA^{-1} = \text{rge}(C_L, I)$  with  $C_L = BA^{-1}$  follows. Clearly,  $C_L$  is uniquely given by  $L$  and does not depend on the particular choice of  $A$  and  $B$ . From  $L = \text{rge}(C_L, I)$  we deduce  $L^\perp = \text{rge}(I, -C_L^T)$  and  $L^* = S_n L^\perp = \text{rge}(C_L^T, I)$ . Further, for every  $p \in \mathbb{R}^n$  we have  $(p, -C_L^T p) \in L^\perp$ , implying

$$\langle p, y^* \rangle - \langle C_L^T p, x^* \rangle = \langle p, y^* - C_L x^* \rangle = 0 \quad \forall (y^*, x^*) \in L$$

and (32) follows. Finally, for every  $(y^*, x^*) \in L$  there is some  $p \in \mathbb{R}^n$  with  $y^* = C_L p$ ,  $x^* = p$  implying (33).  $\square$

**Remark 4.3.** Note that for every  $L \in \mathcal{Z}_n^{\text{reg}}$  there holds  $C_L = \pi_1(Z)\pi_2(Z)^{-1}$  for every  $Z \in \mathcal{M}(L)$ .

In case of SCD regularity we obtain from Proposition 4.2 that

$$\bigcup \mathcal{S}^* F(x, y) = \{(C_L p, p) \mid L \in \mathcal{S}^* F(x, y), p \in \mathbb{R}^n\}$$

and consequently

$$\begin{aligned} \text{scd reg } F(x, y) &= \sup\{\|C_L p\| \mid L \in \mathcal{S}^* F(x, y), p \in \mathbb{R}^n, \|p\| \leq 1\} \\ &= \sup\{\|C_L\| \mid L \in \mathcal{S}^* F(x, y)\}. \end{aligned} \tag{34}$$

**Remark 4.4.** In case of a single-valued, locally Lipschitzian mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , by virtue of Lemma 3.11, SCD regularity of  $F$  around  $\bar{x}$  means that all matrices belonging to the B-subdifferential are nonsingular. This is exactly the so-called BD-regularity property from [29].

By isometry of the mapping  $L \mapsto L^*$  and Proposition 4.2 we obtain readily the following lemma.

**Lemma 4.5.** *The mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is SCD regular around  $(\bar{x}, \bar{y}) \in \text{gph } F$  if and only if*

$$(u, 0) \in \bigcup \mathcal{S}F(x, y) \Rightarrow u = 0. \quad (35)$$

Further,

$$\text{scd reg } F(x, y) = \sup\{\|u\| \mid (u, v) \in \bigcup \mathcal{S}F(x, y), \|v\| \leq 1\} = \sup\{\|C_L\| \mid L \in \mathcal{S}F(x, y)\}.$$

Note that SCD regularity is weaker than the metric regularity of  $F$  around  $(x, y)$ . Indeed, condition (6) for metric regularity of  $F$  near  $(x, y)$  can be equivalently written as

$$(y^*, 0) \in \text{gph } D^*F(x, y) \Rightarrow y^* = 0$$

and  $\bigcup \mathcal{S}^*F(x, y)$  is contained in  $\text{gph } D^*F(x, y)$  by Lemma 3.7. The next example shows that SCD regularity is even strictly weaker than metric regularity, see also Example 6.5 below.

**Example 4.6.** Consider the SCD mapping

$$F(x) := -x + N_{\mathbb{R}_-}(x) = \partial q(x) \quad \text{with } q(x) = -\frac{1}{2}x^2 + \delta_{\mathbb{R}_-}(x)$$

at  $(0, 0)$ . Then

$$D^*F(0, 0)(y^*) = -y^* + \begin{cases} \{0\} & \text{if } y^* < 0 \\ \mathbb{R} & \text{if } y^* = 0 \\ \mathbb{R}_+ & \text{if } y^* > 0 \end{cases}$$

and therefore the only subspaces contained in  $\text{gph } D^*F(0, 0)$  are  $\{(y^*, -y^*) \mid y^* \in \mathbb{R}\}$  and  $\{0\} \times \mathbb{R}$ . Hence  $F$  is SCD regular at  $(0, 0)$ , but  $F$  is not metrically regular near  $(0, 0)$  because of  $0 \in D^*F(0, 0)(1)$ .

**Lemma 4.7.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be SCD regular around  $(x, y) \in \text{gph } F$ . Then  $\text{scd reg } F(x, y) < \infty$ .*

**Proof.** Assume on the contrary that there are sequences  $L_k \in \mathcal{S}^*F(x, y)$  and  $(y_k^*, x_k^*) \in L_k$  such that  $\|y_k^*\| \geq k$  and  $\|x_k^*\| \leq 1$ . By possibly passing to some subsequence we can assume that  $y_k^*/\|y_k^*\|$  converges to some  $y^*$  with  $\|y^*\| = 1$  and  $L_k$  converges in the compact metric space  $\mathcal{L}_n$  to some  $L$ . Then  $(y^*, 0) = \lim_{k \rightarrow \infty} (y_k^*, x_k^*)/\|y_k^*\| \in L$  and  $L \in \mathcal{S}^*F(x, y)$  by Lemma 3.13 contradicting the assumption of SCD regularity.  $\square$

**Proposition 4.8.** *Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is SCD regular around  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then  $F$  is SCD regular around every  $(x, y) \in \text{gph } F$  sufficiently close to  $(\bar{x}, \bar{y})$  and*

$$\limsup_{(x, y) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})} \text{scd reg } F(x, y) \leq \text{scd reg } F(\bar{x}, \bar{y}).$$

**Proof.** By contraposition. If any of the assertions does not hold, we can find some  $\kappa > \text{scd reg } F(\bar{x}, \bar{y})$  and sequences  $(x_k, y_k) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$ ,  $L_k \in \mathcal{S}^*F(x_k, y_k)$  and  $(y_k^*, x_k^*) \in L_k$  with  $\|y_k^*\| \geq \kappa$  and  $\|x_k^*\| \leq 1$ . By possibly passing to some subsequence we can assume that  $(y_k^*, x_k^*)/\|y_k^*\|$  converges to some  $(y^*, x^*)$  and  $L_k$  converges to some  $L$ . Then  $\|y^*\| = 1$ ,  $\|x^*\| \leq \frac{1}{\kappa}$ ,  $(y^*, x^*) \in L$  and  $L \in \mathcal{S}^*F(\bar{x}, \bar{y})$  by Lemma 3.13. Since  $L$  is a subspace, we also have  $(\kappa y^*, \kappa x^*) \in L \subseteq \bigcup \mathcal{S}^*F(\bar{x}, \bar{y})$  implying together with  $\|\kappa x^*\| \leq 1$  the contradiction  $\text{scd reg } F(\bar{x}, \bar{y}) \geq \|\kappa y^*\| = \kappa$ .  $\square$

### 5. On semismooth\* Newton methods for SCD mappings

Consider the inclusion

$$0 \in F(x), \tag{36}$$

where  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ . Assume that  $\bar{x}$  is a reference solution of (36). The idea behind the semismooth\* Newton method [14] for solving (36) is as follows. If  $F$  is semismooth\* at  $(\bar{x}, 0)$  and we are given some point  $(x, y) \in \text{gph } F$  close to  $(\bar{x}, 0)$ , then for every  $(y^*, x^*) \in \text{gph } D^*F(x, y)$  there holds

$$\langle x^*, x - \bar{x} \rangle = \langle y^*, y - 0 \rangle + o(\|(x, y) - (\bar{x}, 0)\| \|(x^*, y^*)\|)$$

by the definition of the semismoothness\* property. We choose now  $n$  pairs  $(y_i^*, x_i^*) \in \text{gph } D^*F(x, y)$ ,  $i = 1, \dots, n$ , compute a solution  $\Delta x$  of the system

$$\langle x_i^*, \Delta x \rangle = -\langle y_i^*, y \rangle, \quad i = 1, \dots, n \tag{37}$$

and expect that  $\|(x + \Delta x) - \bar{x}\| = o(\|(x, y) - (\bar{x}, 0)\|)$ . When dealing with SCD mappings  $F$  we can simplify this procedure by choosing the pairs  $(y_i^*, x_i^*)$  as a basis of some subspace  $L \in \mathcal{S}^*F(x, y)$ , which allows us to weaken the notion of semismoothness\* along the lines of Proposition 2.11.

**Definition 5.1.** We say that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is *SCD semismooth\** at  $(\bar{x}, \bar{y}) \in \text{gph } F$  if  $F$  has the SCD property around  $(\bar{x}, \bar{y})$  and for every  $\varepsilon > 0$  there is some  $\delta > 0$  such that

$$|\langle x^*, x - \bar{x} \rangle - \langle y^*, y - \bar{y} \rangle| \leq \varepsilon \|(x, y) - (\bar{x}, \bar{y})\| \|(x^*, y^*)\| \tag{38}$$

holds for all  $(x, y) \in \text{gph } F \cap \mathcal{B}_\delta(\bar{x}, \bar{y})$  and all  $(y^*, x^*) \in \bigcup \mathcal{S}^*F(x, y)$ .

We say that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is *SCD semismooth\* around*  $(\bar{x}, \bar{y}) \in \text{gph } F$  if there is some neighborhood  $W$  of  $(\bar{x}, \bar{y})$  such that  $F$  is SCD semismooth\* at every  $(x, y) \in \text{gph } F \cap W$ .

The chosen subspace  $L \in \mathcal{S}^*F(x, y)$  should have the property that the resulting system (37) has a unique solution. Taking the pairs  $(y_i^*, x_i^*)$ ,  $i = 1, \dots, n$ , as columns of a  $2n \times n$  matrix  $Z$ , this yields the requirement that  $\pi_2(Z)$  is nonsingular, which in turn is equivalent to  $L \in \mathcal{L}_n^{\text{reg}}$  by Proposition 4.2. Comparing (32) with (37), we see that  $\Delta x = -C_L^T y$  is a solution to (37) in this case.

We are now in the position to describe the iteration step of the SCD variant of the semismooth\* Newton method introduced in [14]. Assume we are given some iterate  $x^{(k)}$ . Since we cannot expect in general that  $F(x^{(k)}) \neq \emptyset$  or that  $0$  is close to  $F(x^{(k)})$ , even if  $x^{(k)}$  is close to a solution  $\bar{x}$ , we first perform some preparatory step which yields  $(\hat{x}^{(k)}, \hat{y}^{(k)}) \in \text{gph } F$  as an approximate projection of  $(x^{(k)}, 0)$  onto  $\text{gph } F$ . We require that

$$\|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (\bar{x}, 0)\| \leq \eta \|x^{(k)} - \bar{x}\| \tag{39}$$

for some constant  $\eta > 0$ . E.g., if

$$\|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (x^{(k)}, 0)\| \leq \beta \text{dist}((x^{(k)}, 0), \text{gph } F)$$

holds with some  $\beta \geq 1$ , then

$$\begin{aligned} \|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (\bar{x}, 0)\| &\leq \|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (x^{(k)}, 0)\| + \|(x^{(k)}, 0) - (\bar{x}, 0)\| \\ &\leq \beta \text{dist}((x^{(k)}, 0), \text{gph } F) + \|(x^{(k)}, 0) - (\bar{x}, 0)\| \leq (\beta + 1) \|(x^{(k)}, 0) - (\bar{x}, 0)\| \end{aligned}$$

and (39) holds with  $\eta = \beta + 1$ . Further we require that  $\mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)}) \cap \mathcal{Z}_n^{\text{reg}} \neq \emptyset$  and compute the new iterate as  $x^{(k+1)} = \hat{x}^{(k)} - C_L^T \hat{y}^{(k)}$  for some  $L \in \mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)}) \cap \mathcal{Z}_n^{\text{reg}}$ . In fact, in a numerical implementation we will not calculate the matrix  $C_L$ , but two  $n \times n$  matrices  $A, B$  such that  $L = \text{rge}(B^T, A^T)$ , compute  $\Delta x^{(k)}$  as a solution of the system  $A\Delta x = -B\hat{y}^{(k)}$  and set  $x^{(k+1)} = \hat{x}^{(k)} + \Delta x^{(k)}$ .

This leads to the following conceptual algorithm.

**Algorithm 1** (SCD semismooth\* Newton-type method for inclusions).

1. Choose a starting point  $x^{(0)}$ , set the iteration counter  $k := 0$ .
2. If  $0 \in F(x^{(k)})$ , stop the algorithm.
3. **Approximation step:** Compute

$$(\hat{x}^{(k)}, \hat{y}^{(k)}) \in \text{gph } F$$

satisfying (39) such that  $\mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)}) \cap \mathcal{Z}_n^{\text{reg}} \neq \emptyset$ .

4. **Newton step:** Select  $n \times n$  matrices  $A^{(k)}, B^{(k)}$  with  $L^{(k)} := \text{rge}(B^{(k)T}, A^{(k)T}) \in \mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)}) \cap \mathcal{Z}_n^{\text{reg}}$ , calculate the Newton direction  $\Delta x^{(k)}$  as a solution of the linear system  $A^{(k)}\Delta x = -B^{(k)}\hat{y}^{(k)}$  and obtain the new iterate via  $x^{(k+1)} = \hat{x}^{(k)} + \Delta x^{(k)}$ .
5. Set  $k := k + 1$  and go to 2.

We have  $\Delta x^{(k)} = -C_{L^{(k)}}^T \hat{y}^{(k)}$  and therefore  $(\Delta x^{(k)}, -\hat{y}^{(k)}) \in -L^{(k)*} = L^{(k)*} \in \mathcal{S}F(\hat{x}^{(k)}, \hat{y}^{(k)})$  by Proposition 4.2. Thus, alternatively we can perform the Newton step also in the following way:

4. **Newton step:** Select  $n \times n$  matrices  $A^{(k)}, B^{(k)}$  with  $\text{rge}(B^{(k)}, A^{(k)}) \in \mathcal{S}F(\hat{x}^{(k)}, \hat{y}^{(k)}) \cap \mathcal{Z}_n^{\text{reg}}$ , compute a solution  $p$  of the linear system  $A^{(k)}p = -\hat{y}^{(k)}$  and compute the new iterate  $x^{(k+1)} = \hat{x}^{(k)} + \Delta x^{(k)}$  with Newton direction  $\Delta x^{(k)} = B^{(k)}p$ .

**Remark 5.2.** Note that  $-\hat{y}^{(k)} \in D^\#F(\hat{x}^{(k)}, \hat{y}^{(k)})(\Delta x^{(k)})$  but we do not necessarily have  $-\hat{y}^{(k)} \in DF(\hat{x}^{(k)}, \hat{y}^{(k)})(\Delta x^{(k)})$  as it is the case in Newton methods based on the graphical derivative, cf. [7,17,25].

Which possibility for calculating the Newton direction is actually chosen, depends on the availability of the respective derivative. Let us analyze the two alternatives for the special case when  $F$  is single-valued and continuously differentiable at  $x^{(k)}$ . In Algorithm 1, the matrices  $A^{(k)}, B^{(k)}$  with  $\text{rge}(B^{(k)T}, A^{(k)T}) \in \mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)})$  fulfill  $A^{(k)T} = \nabla F(\hat{x}^{(k)})^T B^{(k)T}$  and thus the Newton direction is computed by solving the linear system  $(B^{(k)}\nabla F(\hat{x}^{(k)}))\Delta x = -B^{(k)}\hat{y}^{(k)}$ . On the other hand, given  $A^{(k)}, B^{(k)}$  with  $\text{rge}(B^{(k)}, A^{(k)}) \in \mathcal{S}F(\hat{x}^{(k)}, \hat{y}^{(k)})$ , we have  $A^{(k)} = \nabla F(\hat{x}^{(k)})B^{(k)}$  and in this case the Newton direction is computed via  $\Delta x^{(k)} = B^{(k)}p = -B^{(k)}(\nabla F(\hat{x}^{(k)})B^{(k)})^{-1}\hat{y}^{(k)}$ . The structure of the second approach resembles the adjoint system method known from PDE-constrained optimization and optimal control.

We now consider convergence of Algorithm 1.

**Proposition 5.3.** Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is SCD semismooth\* at  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then for every  $\varepsilon > 0$  there is some  $\delta > 0$  such that the inequality

$$\|x - C_L^T(y - \bar{y}) - \bar{x}\| \leq \varepsilon \sqrt{n(1 + \|C_L\|^2)} \|(x, y) - (\bar{x}, \bar{y})\| \tag{40}$$

holds for every  $(x, y) \in \text{gph } F \cap \mathcal{B}_\delta(\bar{x}, \bar{y})$  and every  $L \in \mathcal{S}^*F(x, y) \cap \mathcal{Z}_n^{\text{reg}}$ .



**Proof.** Pick  $\varepsilon > 0$  and choose  $\delta > 0$  such that (38) holds. Now consider any  $(x, y) \in \text{gph } F \cap \mathcal{B}_\delta(\bar{x}, \bar{y})$  and any  $L \in \mathcal{S}^*F(x, y) \cap \mathcal{X}_n^{\text{reg}}$ . By Proposition 4.2 we have  $L = \text{rge}(C_L, I)$  and therefore  $(C_L e_i, e_i) \in L$ ,  $i = 1, \dots, n$ , where  $e_i$  denotes the  $i$ -th unit vector. From (38) we obtain

$$\begin{aligned} |\langle e_i, \bar{x} - x \rangle - \langle C_L e_i, \bar{y} - y \rangle| &= |\langle e_i, x - C_L^T(y - \bar{y}) - \bar{x} \rangle| \leq \varepsilon \|(e_i, C_L e_i)\| \|(x, y) - (\bar{x}, \bar{y})\| \\ &\leq \varepsilon \sqrt{1 + \|C_L\|^2} \|(x, y) - (\bar{x}, \bar{y})\| \end{aligned}$$

and

$$\|x - C_L^T(y - \bar{y}) - \bar{x}\| \leq \varepsilon \sqrt{n(1 + \|C_L\|^2)} \|(x, y) - (\bar{x}, \bar{y})\|$$

follows.  $\square$

Given  $\eta, \kappa > 0$ , we now define for  $x \in \mathbb{R}^n$  the set

$$\mathcal{G}_{F, \bar{x}}^{\eta, \kappa}(x) := \{(\hat{x}, \hat{y}, L) \mid (\hat{x}, \hat{y}) \in \text{gph } F, \|(\hat{x}, \hat{y}) - (\bar{x}, 0)\| \leq \eta \|x - \bar{x}\|, L \in \mathcal{S}^*F(\hat{x}, \hat{y}) \cap \mathcal{X}_n^{\text{reg}}, \|C_L\| \leq \kappa\}.$$

**Theorem 5.4.** Assume that  $F$  is SCD semismooth\* at  $(\bar{x}, 0) \in \text{gph } F$  and assume that there are  $\eta, \kappa > 0$  such that for every  $x \notin F^{-1}(0)$  sufficiently close to  $\bar{x}$  we have  $\mathcal{G}_{F, \bar{x}}^{L, \kappa}(x) \neq \emptyset$ . Then there exists some  $\delta > 0$  such that for every starting point  $x^{(0)} \in \mathcal{B}_\delta(\bar{x})$  Algorithm 1 either stops after finitely many iterations at a solution or produces a sequence  $x^{(k)}$  which converges superlinearly to  $\bar{x}$ , provided we choose in every iteration  $(\hat{x}^{(k)}, \hat{y}^{(k)}, L^{(k)}) \in \mathcal{G}_{F, \bar{x}}^{\eta, \kappa}(x^{(k)})$ .

**Proof.** Using Proposition 5.3 with  $\bar{y} = 0$ , we can find some  $\bar{\delta} > 0$  such that (40) holds with  $\varepsilon = \frac{1}{2\eta\sqrt{n(1+\kappa^2)}}$  for all  $(x, y) \in \text{gph } F \cap \mathcal{B}_{\bar{\delta}}(\bar{x}, 0)$  and all  $L \in \mathcal{S}^*F(x, y) \cap \mathcal{X}_n^{\text{reg}}$ . Set  $\delta := \bar{\delta}/\eta$  and consider an iterate  $x^{(k)} \in \mathcal{B}_\delta(\bar{x}) \notin F^{-1}(0)$ . Then

$$\|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (\bar{x}, 0)\| \leq \eta \|x^{(k)} - \bar{x}\| \leq \bar{\delta}$$

and consequently

$$\|x^{(k+1)} - \bar{x}\| \leq \frac{1}{2\eta\sqrt{n(1+\kappa^2)}} \sqrt{n(1+\kappa^2)} \|(\hat{x}^{(k)}, \hat{y}^{(k)}) - (\bar{x}, 0)\| \leq \frac{1}{2} \|x^{(k)} - \bar{x}\|$$

by Proposition 5.3. It follows that for every starting point  $x^{(0)} \in \mathcal{B}_\delta(\bar{x})$  Algorithm 1 either stops after finitely many iterations with a solution or produces a sequence  $x^{(k)}$  converging to  $\bar{x}$ . The superlinear convergence of the sequence  $x^{(k)}$  is now an easy consequence of Proposition 5.3.  $\square$

So far Algorithm 1 is only a straightforward adaption of the semismooth\* Newton method from [14] to SCD mappings. However, in [14] the semismooth\* Newton method was only guaranteed to converge under the assumption of strong metric regularity, whereas we will now prove that for its SCD variant a less restrictive condition is sufficient.

**Proposition 5.5.** Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be SCD regular around  $(\bar{x}, 0) \in \text{gph } F$ . Then for every  $\eta > 0$  and every  $\kappa > \text{scd reg } F(\bar{x}, 0)$  there is a neighborhood  $U$  of  $\bar{x}$  such that for every  $x \in U$  the set  $\mathcal{G}_{F, \bar{x}}^{\eta, \kappa}(x)$  is nonempty and amounts to

$$\mathcal{G}_{F, \bar{x}}^{\eta, \kappa}(x) = \{(\hat{x}, \hat{y}, L) \mid (\hat{x}, \hat{y}) \in \text{gph } F, \|(\hat{x}, \hat{y}) - (\bar{x}, 0)\| \leq \eta \|x - \bar{x}\|, L \in \mathcal{S}^*F(\hat{x}, \hat{y})\}. \tag{41}$$

**Proof.** By Proposition 4.8 we can find some positive radius  $\rho$  such that  $F$  is SCD regular around  $(x, y)$  with modulus  $\text{scd reg } F(x, y) \leq \kappa$  for every  $(x, y) \in \text{gph } F \cap \mathcal{B}_\rho(\bar{x}, 0)$ . By taking  $U := \mathcal{B}_{\rho/\eta}(\bar{x})$ , for every  $x \in U$  and every  $(\hat{x}, \hat{y}) \in \text{gph } F$  with  $\|(\hat{x}, \hat{y}) - (\bar{x}, 0)\| \leq \eta\|x - \bar{x}\|$  we have  $(\hat{x}, \hat{y}) \in \mathcal{B}_\rho(\bar{x}, 0)$ . By (34) we obtain that  $\|C_L\| \leq \kappa$  whenever  $L \in \mathcal{S}^*F(\hat{x}, \hat{y})$  and the assertion follows.  $\square$

Since in (41) the right hand side does not depend on  $\kappa$ , we obtain the following corollary of Theorem 5.4.

**Corollary 5.6.** *Assume that  $F$  is SCD semismooth\* at  $(\bar{x}, 0) \in \text{gph } F$  and SCD regular around  $(\bar{x}, 0)$ . Then for every  $\eta > 0$  there is a neighborhood  $U$  of  $\bar{x}$  such that for every starting point  $x^{(0)} \in U$  Algorithm 1 is well-defined and either stops after finitely many iterations at a solution of (36) or produces a sequence  $x^{(k)}$  converging superlinearly to  $\bar{x}$  for any choice of  $(\hat{x}^{(k)}, \hat{y}^{(k)})$  satisfying (39) and any  $L^{(k)} \in \mathcal{S}^*F(\hat{x}^{(k)}, \hat{y}^{(k)})$ .*

**Remark 5.7.** Note that Corollary 5.6 guarantees not only locally superlinear convergence, but also that the method is locally well-defined, which is an advantage in comparison with the Josephy-Newton method from [18]. In Theorem 6.2 below we will show that, under the assumptions of Corollary 5.6, the mapping  $F$  is strongly metrically subregular at  $(\bar{x}, 0)$ . By [5, Theorem 6.1], in such a case the convergence of the Josephy-Newton method is also locally superlinear, provided the method is well-defined. This, however, need not be the case as illustrated in [14, Example 5.13], where the assumptions of Corollary 5.6 are fulfilled, the semismooth\* Newton method works well, but the Josephy-Newton method collapses.

## 6. Strong metric subregularity on a neighborhood

We first present a characterization of strong metric subregularity on a neighborhood, cf. Definition 2.8, by means of the outer limiting graphical derivative defined in Definition 3.8.

**Theorem 6.1.** *Consider a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then  $F$  is strongly metrically subregular around  $(\bar{x}, \bar{y})$  if and only if the condition*

$$0 \in D^\sharp F(\bar{x}, \bar{y})(u) \Rightarrow u = 0 \quad (42)$$

holds and in this case one has

$$\text{l-subreg } F(\bar{x}, \bar{y}) = \sup\{\|u\| \mid (u, v) \in \text{gph } D^\sharp F(\bar{x}, \bar{y}), \|v\| \leq 1\}. \quad (43)$$

**Proof.** We prove the “if”-part by contraposition. Assume that (42) holds but  $F$  is not strongly metrically subregular around  $(\bar{x}, \bar{y})$ . Then we can find a sequence  $(x_k, y_k) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$  such that either  $F$  is not strongly metrically subregular at  $(x_k, y_k)$  for infinitely many  $k$  or  $\limsup_{k \rightarrow \infty} \text{subreg } F(x_k, y_k) = \infty$ . After possibly passing to some subsequence and taking into account Theorem 2.7, in both cases there is a sequence  $\kappa_k \rightarrow \infty$  and  $(u_k, v_k) \in \text{gph } DF(x_k, y_k)$  with  $\|v_k\| \leq 1$  such that  $\|u_k\| > \kappa_k \|v_k\|$ . Defining  $(\tilde{u}_k, \tilde{v}_k) := (u_k, v_k) / \|u_k\| \in \text{gph } DF(x_k, y_k)$ , we have  $\|\tilde{v}_k\| \leq 1/\kappa_k$  implying  $\lim_{k \rightarrow \infty} \tilde{v}_k = 0$ . By possibly passing to a subsequence once more,  $\tilde{u}_k$  converges to some  $u$  with  $\|u\| = 1$  and from the definition of  $D^\sharp F(\bar{x}, \bar{y})$  we obtain  $(u, 0) \in \text{gph } D^\sharp F(\bar{x}, \bar{y})$  contradicting (42). This proves the “if”-part.

In order to show the “only if”-part assume that (42) does not hold, so that there is some  $u \neq 0$  with  $(u, 0) \in \text{gph } D^\sharp F(\bar{x}, \bar{y})$ . By definition of  $D^\sharp F(\bar{x}, \bar{y})$  there are sequences  $(x_k, y_k) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$  and  $(u_k, v_k) \rightarrow (u, 0)$  with  $(u_k, v_k) \in \text{gph } DF(x_k, y_k)$ . If  $F$  is not strongly metrically subregular at  $(x_k, y_k)$  for infinitely many  $k$ , then it is not strongly metrically subregular around  $(\bar{x}, \bar{y})$  by definition. On the other hand, if  $F$  is strongly metrically subregular at  $(x_k, y_k)$  then  $v_k \neq 0$  and  $\text{subreg } F(x_k, y_k) \geq \|u_k\| / \|v_k\|$ , which follows from Theorem 2.7. Hence,  $\limsup_{k \rightarrow \infty} \text{subreg } F(x_k, y_k) = \infty$  and  $F$  is again not strongly metrically subregular around  $(\bar{x}, \bar{y})$ . This proves the “only if”-part. There remains to show (43). By definition we have

$$\text{l-subreg } F(\bar{x}, \bar{y}) = \limsup_{(x,y) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})} \sup\{\|u\| \mid (u, v) \in \text{gph } DF(x, y), \|v\| \leq 1\}$$

and therefore there are sequences  $(x_k, y_k) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$  and  $(u_k, v_k) \in \text{gph } DF(x_k, y_k)$  with  $\|v_k\| \leq 1$  and  $\|u_k\| \rightarrow \text{l-subreg } F(\bar{x}, \bar{y}) < \infty$ . By possibly passing to a subsequence,  $(u_k, v_k)$  converges to some  $(u, v)$  with  $\|v\| \leq 1$ . By definition of  $D^\sharp F$  we have  $(u, v) \in \text{gph } D^\sharp F(\bar{x}, \bar{y})$  and  $\text{l-subreg } F(\bar{x}, \bar{y}) = \|u\| \leq \xi := \sup\{\|u\| \mid \exists v : (u, v) \in \text{gph } D^\sharp F(\bar{x}, \bar{y}), \|v\| \leq 1\}$  follows. Next consider a sequence  $(u_k, v_k) \in \text{gph } D^\sharp F(\bar{x}, \bar{y})$  with  $\|v_k\| \leq 1$  and  $\|u_k\| \rightarrow \xi$ . Then for every  $k$  there are  $(x_k, y_k) \in \text{gph } F$  and  $(u'_k, v'_k) \in \text{gph } DF(x_k, y_k)$  such that  $\|(x_k, y_k) - (\bar{x}, \bar{y})\| \leq \frac{1}{k}$  and  $\|(u'_k, v'_k) - (u_k, v_k)\| \leq \frac{1}{k}$  and

$$\text{subreg } F(x_k, y_k) \geq \xi_k := \begin{cases} \frac{\|u'_k\|}{\|v'_k\|} & \text{if } \|v'_k\| > 1 \\ \|u'_k\| & \text{if } \|v'_k\| \leq 1 \end{cases}$$

follows. In case when  $\|v'_k\| > 1$  we have

$$\left| \frac{\|u'_k\|}{\|v'_k\|} - \|u'_k\| \right| = \frac{\|v'_k\| - 1}{\|v'_k\|} \|u'_k\| \leq \frac{1}{k+1} \|u'_k\|$$

and  $\lim_{k \rightarrow \infty} \xi_k = \xi$  follows. Hence  $\text{l-subreg } F(\bar{x}, \bar{y}) \geq \limsup_{k \rightarrow \infty} \xi_k = \xi$  and relation (43) is established.  $\square$

From Lemma 3.10 together with Lemma 4.5, we may conclude that strong metric subregularity around  $(\bar{x}, \bar{y}) \in \text{gph } F$  implies SCD regularity around  $(\bar{x}, \bar{y})$  and that  $\text{l-subreg } F(\bar{x}, \bar{y}) \geq \text{scd reg } F(\bar{x}, \bar{y})$ . Next we show that, conversely, SCD regularity in conjunction with SCD semismoothness\* provides a sufficient condition for strong metric subregularity.

**Theorem 6.2.** *Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is SCD regular around  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then for every  $\kappa > \text{scd reg } F(\bar{x}, \bar{y})$  there is a neighborhood  $W$  of  $(\bar{x}, \bar{y})$  such that  $F$  is strongly metrically subregular with modulus  $\text{subreg } F(x, y) < \kappa$  at every point  $(x, y) \in \text{gph } F \cap W$  where  $F$  is SCD semismooth\*.*

**Proof.** Fixing  $\kappa > \tilde{\kappa} > \text{scd reg } F(\bar{x}, \bar{y})$ , by Proposition 4.8 there is an open neighborhood  $W$  of  $(\bar{x}, \bar{y})$  such that for every  $(x, y) \in \text{gph } F \cap W$  the mapping  $F$  is SCD regular around  $(x, y)$  with modulus  $\text{scd reg } F(x, y) \leq \tilde{\kappa}$ . Consider  $(\tilde{x}, \tilde{y}) \in \text{gph } F \cap W$  where  $F$  is SCD semismooth\*. Assume now that  $F$  is not metrically subregular at  $(\tilde{x}, \tilde{y})$  or that  $\text{subreg } F(\tilde{x}, \tilde{y}) > \tilde{\kappa}$ . Then there is some  $\kappa' > \tilde{\kappa}$  and a sequence  $x_k$  converging to  $\tilde{x}$  such that  $\text{dist}(x_k, F^{-1}(\tilde{y})) > \kappa' \text{dist}(\tilde{y}, F(x_k)) \forall k$ . Consider  $y_k \in F(x_k)$  with  $\text{dist}(\tilde{y}, F(x_k)) = \|y_k - \tilde{y}\|$ . Then  $y_k$  converges to  $\tilde{y}$  and for all  $k$  sufficiently large we have  $(x_k, y_k) \in \text{gph } F \cap W$ . Pick  $L_k \in \mathcal{S}^*F(x_k, y_k)$ . Using Proposition 5.3 and (34) we have that  $\|C_{L_k}\| \leq \tilde{\kappa}$  and

$$\|x_k - C_{L_k}^T(y_k - \tilde{y}) - \tilde{x}\| \leq \frac{1}{k} \sqrt{n(1 + \|C_{L_k}\|^2)} \|(x_k - \tilde{x}, y_k - \tilde{y})\| \leq \frac{1}{k} \sqrt{n(1 + \tilde{\kappa}^2)} (\|x_k - \tilde{x}\| + \|y_k - \tilde{y}\|)$$

implying

$$(1 - \alpha_k) \|x_k - \tilde{x}\| \leq \|C_{L_k}(y_k - \tilde{y})\| + \alpha_k \|y_k - \tilde{y}\| \leq (\tilde{\kappa} + \alpha_k) \|y_k - \tilde{y}\|,$$

where  $\alpha_k := \frac{1}{k} \sqrt{n(1 + \tilde{\kappa}^2)}$ . Since  $\alpha_k \rightarrow 0$  as  $k \rightarrow \infty$ , we have  $(\tilde{\kappa} + \alpha_k)/(1 - \alpha_k) < \kappa'$  for all  $k$  sufficiently large and therefore  $\|x_k - \tilde{x}\| < \kappa' \|y_k - \tilde{y}\|$  in contrary to our assumption. This shows that  $F$  is metrically subregular at  $(\tilde{x}, \tilde{y})$  and  $\text{subreg } F(\tilde{x}, \tilde{y}) \leq \tilde{\kappa} < \kappa$ . Further  $\tilde{x}$  must be an isolated point in  $F^{-1}(\tilde{y})$ . Assume on the contrary that there is a sequence  $x_k \in F^{-1}(\tilde{y})$  converging to  $\tilde{x}$ . Taking  $L_k \in \mathcal{S}^*F(x_k, \tilde{y})$  and applying Proposition 5.3 with  $\varepsilon = 1/(2\sqrt{n(1 + \tilde{\kappa}^2)})$ , we obtain for all  $k$  sufficiently large

$$\|x_k - \tilde{x}\| = \|x_k - C_{L_k}^T(\tilde{y} - \bar{y}) - \tilde{x}\| \leq \frac{1}{2}\|(x_k - \tilde{x}, \tilde{y} - \bar{y})\| = \frac{1}{2}\|x_k - \tilde{x}\|,$$

a contradiction. This shows that  $F$  is even strongly metrically subregular at  $(\bar{x}, \bar{y})$ .  $\square$

**Remark 6.3.** In the special case of a single-valued locally Lipschitzian mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  the statement of Theorem 6.2 can be derived also from [15, Proposition 1].

For semismooth\* mappings we arrive thus at the following equivalence.

**Corollary 6.4.** *Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is SCD semismooth\* around  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then  $F$  is strongly metrically subregular around  $(\bar{x}, \bar{y})$  if and only if  $F$  is SCD regular around  $(\bar{x}, \bar{y})$  and in this case one has  $\text{l-subreg } F(\bar{x}, \bar{y}) = \text{scd reg } F(\bar{x}, \bar{y})$ .*

**Example 6.5.** Consider the mapping  $F := \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$  given by

$$F(x_1, x_2) := \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} + h(x) + N_C(x_1, x_2),$$

where  $C := \{(x_1, x_2) \mid -\frac{1}{2}x_1 \leq x_2 \leq \frac{1}{2}x_1\}$  is a convex polyhedral cone and  $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is any continuously differentiable mapping satisfying  $h(0, 0) = (0, 0)$ ,  $\nabla h(0, 0) = 0$ . As the reference point we take  $(\bar{x}, \bar{y}) = ((0, 0), (0, 0))$ . The mapping  $N_C$  is a polyhedral mapping and therefore semismooth\* at any point of its graph by [14]. Further,  $N_C(x) = \partial\delta_C(x)$  is an SCD mapping by Corollary 3.28. Thus,  $F$  is both an SCD mapping around and semismooth\* at any point of its graph, because it differs from  $N_C$  by a continuously differentiable mapping. Now let us calculate  $\mathcal{S}^*F(\bar{x}, \bar{y})$ . The critical cone  $\mathcal{H}_C(\bar{x}, \bar{y})$  amounts to  $C$  and has therefore the 4 faces  $C$ ,  $\{(u, \frac{1}{2}u) \mid u \geq 0\}$ ,  $\{(u, -\frac{1}{2}u) \mid u \geq 0\}$  and  $\{(0, 0)\}$ . By using Example 3.29 we conclude that  $\mathcal{S}^*N_C(0, 0)$  consists of the 4 subspaces  $L_1 := \mathbb{R}^2 \times \{(0, 0)\}$ ,  $L_2 := \{((u, \frac{1}{2}u), (-\frac{1}{2}v, v)) \mid (u, v) \in \mathbb{R}^2\}$ ,  $L_3 := \{((u, -\frac{1}{2}u), (\frac{1}{2}v, v)) \mid (u, v) \in \mathbb{R}^2\}$  and  $L_4 := \{(0, 0)\} \times \mathbb{R}^2$  and Proposition 3.15 tells us that  $\mathcal{S}^*F(\bar{x}, \bar{y}) = \{TL_1, TL_2, TL_3, TL_4\}$  where

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$$

Straightforward calculations yield

$$TL_1 = \{((u, v), (u, -v)) \mid (u, v) \in \mathbb{R}^2\}, TL_2 = ((u, \frac{1}{2}u), (u - \frac{1}{2}v, -\frac{1}{2}u + v)) \mid (u, v) \in \mathbb{R}^2\},$$

$$TL_3 = \{((u, -\frac{1}{2}u), (u + \frac{1}{2}v, \frac{1}{2}u + v)) \mid (u, v) \in \mathbb{R}^2\}, TL_4 = \{((0, 0), (u, v)) \mid (u, v) \in \mathbb{R}^2\}.$$

Now it easily follows that  $F$  is SCD regular around  $(\bar{x}, \bar{y})$  with

$$C_{TL_1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, C_{TL_2} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, C_{TL_3} = \begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix}, C_{TL_4} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and  $\|C_{TL_1}\| = 1$ ,  $\|C_{TL_2}\| = \|C_{TL_3}\| = \frac{5}{3}$ ,  $\|C_{TL_4}\| = 0$ . Hence, by virtue of Theorem 6.2,  $F$  is strongly metrically subregular around  $(\bar{x}, \bar{y})$  with modulus  $\text{l-subreg } F(\bar{x}, \bar{y}) = \frac{5}{3}$ .

To illustrate this result, we explicitly compute  $F^{-1}$  in case  $h = 0$ . One obtains that

$$F^{-1}(y) = \begin{cases} \{z_1(y)\} & \text{if } -\frac{1}{2}y_1 + y_2 > 0, 2y_1 + y_2 \geq 0, \\ \{z_1(y), z_2(y), z_3(y)\} & \text{if } -\frac{1}{2}y_1 + y_2 \leq 0, -\frac{1}{2}y_1 - y_2 \leq 0, \\ \{z_3(y)\} & \text{if } -\frac{1}{2}y_1 - y_2 > 0, 2y_1 - y_2 \geq 0, \\ \{z_4(y)\} & \text{if } 2y_1 + y_2 \leq 0, 2y_1 - y_2 \leq 0 \}, \end{cases} \tag{44}$$

with the mappings  $z_i(y)$ ,  $i = 1, \dots, 4$ , in (44) specified via

$$z_1(y) := \left(\frac{4}{3}y_1 + \frac{2}{3}y_2, \frac{2}{3}y_1 + \frac{1}{3}y_2\right), \quad z_2(y) := (y_1, -y_2),$$

$$z_3(y) := \left(\frac{4}{3}y_1 - \frac{2}{3}y_2, -\frac{2}{3}y_1 + \frac{1}{3}y_2\right), \quad z_4(y) := (0, 0).$$

We see that  $F^{-1}$  has the isolated calmness property at every point of its graph close to  $(0, 0)$ , but it is not single-valued.

### 7. On strong metric regularity

Our results on strong metric regularity pertain again mappings  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and are partly expressed in terms of certain bases for the subspaces  $L \in \mathcal{S}F(\bar{x}, \bar{y})$ .

Given a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  which is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y}) \in \text{gph } F$  with transformation mapping  $\Phi$  according to Definition 3.16, we denote by  $\bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  the collection of all  $2n \times n$  matrices  $Z$  such that  $\text{rge } Z \in \mathcal{S}F(\bar{x}, \bar{y})$  and  $\pi_1(\nabla\Phi(\bar{x}, \bar{y})Z) = I$ .

Note that by Proposition 3.17 for every  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  there exists a unique  $Z \in \bar{\nabla}^\Phi F(\bar{x}, \bar{y}) \cap \mathcal{M}(L)$ . Since for every Lipschitzian mapping  $f : U \rightarrow \mathbb{R}^n$ ,  $U \subset \mathbb{R}^n$  open, the B-subdifferential  $\bar{\nabla}f(u)$  is compact for every  $u \in U$ , we conclude from (24a) that  $\bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  is compact as well.

The next statement provides us with an upper approximation of the graphs of the strict derivative and the limiting coderivative, respectively.

**Proposition 7.1.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be graphically Lipschitzian at  $(\bar{x}, \bar{y}) \in \text{gph } F$  with transformation mapping  $\Phi$ . Then*

$$\text{gph } D_*F(\bar{x}, \bar{y}) \subseteq \bigcup_{Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})} \text{rge } Z, \tag{45}$$

$$\text{gph } D^*F(\bar{x}, \bar{y}) \subseteq \bigcup_{Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})} (\text{rge } Z)^* \tag{46}$$

**Proof.** According to Definition 3.16 consider the open neighborhoods  $W$  of  $(\bar{x}, \bar{y})$ ,  $U$  of  $\bar{w}$  and the Lipschitzian mapping  $f : U \rightarrow \mathbb{R}^n$  with  $\Phi(\text{gph } F \cap W) = \text{gph } f$ , where  $\bar{w} = \pi_1(\Phi(\bar{x}, \bar{y}))$ . Consider  $(u, v) \in \text{gph } D_*F(\bar{x}, \bar{y})$  together with sequences  $(x_k^1, y_k^1) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$ ,  $(x_k^2, y_k^2) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$  and  $t_k \downarrow 0$  such that  $(u, v) = \lim_{k \rightarrow \infty} (x_k^2 - x_k^1, y_k^2 - y_k^1)/t_k$ . For each  $k$  let  $w_k^i$ ,  $i = 1, 2$  be given by  $w_k^i = \pi_1(\Phi(x_k^i, y_k^i))$ . Then

$$\begin{aligned} (w_k^2 - w_k^1, f(w_k^2) - f(w_k^1)) &= \Phi(x_k^2, y_k^2) - \Phi(x_k^1, y_k^1) = \nabla\Phi(\bar{x}, \bar{y})(x_k^2 - x_k^1, y_k^2 - y_k^1) + o(\|(x_k^2 - x_k^1, y_k^2 - y_k^1)\|) \\ &= \nabla\Phi(\bar{x}, \bar{y})(x_k^2 - x_k^1, y_k^2 - y_k^1) + o(t_k) \end{aligned}$$

implying that

$$\lim_{k \rightarrow \infty} \frac{(w_k^2 - w_k^1, f(w_k^2) - f(w_k^1))}{t_k} = \lim_{k \rightarrow \infty} \nabla \Phi(\bar{x}, \bar{y}) \frac{(x_k^2 - x_k^1, y_k^2 - y_k^1)}{t_k} = \nabla \Phi(\bar{x})(u, v) \in \text{gph } D_* f(\bar{w}).$$

Hence  $\pi_2(\nabla \Phi(\bar{x})(u, v)) \in D_* f(\bar{w})(\pi_1(\nabla \Phi(\bar{x})(u, v)))$  and by [34, Theorem 9.62] there is some  $B \in \text{conv } \bar{\nabla} f(\bar{w})$  satisfying  $\pi_2(\nabla \Phi(\bar{x})(u, v)) = B\pi_1(\nabla \Phi(\bar{x})(u, v))$  which is the same as  $\nabla \Phi(\bar{x})(u, v) \in \text{rge}(I, B)$ .  $B$  can be expressed as a convex combination  $\sum_{i=1}^N \alpha_i B_i$  with  $B_i \in \bar{\nabla} f(\bar{w})$ ,  $\alpha_i \geq 0$ ,  $\sum_{i=1}^N \alpha_i = 1$  and therefore

$$(u, v) \in \nabla \Phi(\bar{x}, \bar{y})^{-1} \text{rge}(I, B) = \Phi(\bar{x}, \bar{y})^{-1} \text{rge} \left[ \sum_{i=1}^N \alpha_i \begin{pmatrix} I \\ B_i \end{pmatrix} \right] = \text{rge} \left[ \sum_{i=1}^N \alpha_i \Phi(\bar{x}, \bar{y})^{-1} \begin{pmatrix} I \\ B_i \end{pmatrix} \right].$$

Denoting  $Z_i := \Phi(\bar{x}, \bar{y})^{-1} \begin{pmatrix} I \\ B_i \end{pmatrix}$  we have  $\text{rge } Z_i \in \mathcal{S}F(\bar{x}, \bar{y})$  by (24a) and  $\pi_1(\Phi(\bar{x}, \bar{y})Z_i) = I$  yielding  $Z_i \in \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$ . Thus  $(u, v) \in \text{rge } Z$  with  $Z = \sum_{i=1}^N \alpha_i Z_i \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  verifying (45).

Now consider  $(y^*, x^*) \in \text{gph } D^* F(\bar{x}, \bar{y})$  which is the same as

$$S_n^T(y^*, x^*) \in N_{\text{gph } F}(\bar{x}, \bar{y}) = \nabla \Phi(\bar{x}, \bar{y})^T N_{\text{gph } f}(\bar{w}, f(\bar{w})),$$

where the second equality follows from [34, Exercise 6.7]. Hence  $z^* := S_n \nabla \Phi(\bar{x}, \bar{y})^{-T} S_n^T(y^*, x^*) \in S_n N_{\text{gph } f}(\bar{w}, f(\bar{w})) = \text{gph } D^* f(\bar{w})$  implying  $\pi_2(z^*) \in D^* f(\bar{w})(\pi_1(z^*))$ . By [34, Theorem 9.62] there is some  $B \in \text{conv } \bar{\nabla} f(\bar{w})$  such that  $\pi_2(z^*) = B^T \pi_1(z^*)$  which is the same as  $z^* \in \text{rge}(I, B^T)$  and

$$(y^*, x^*) \in \text{rge} \left[ S_n \nabla \Phi(\bar{x}, \bar{y})^T S_n^T \begin{pmatrix} I \\ B^T \end{pmatrix} \right]$$

follows. Taking into account that

$$\text{rge} \left[ S_n \nabla \Phi(\bar{x}, \bar{y})^T S_n^T \begin{pmatrix} I \\ B^T \end{pmatrix} \right]^\perp = \text{rge} \left[ S_n \nabla \Phi(\bar{x}, \bar{y})^{-1} S_n^T \begin{pmatrix} B \\ -I \end{pmatrix} \right] = \text{rge} \left[ -S_n \nabla \Phi(\bar{x}, \bar{y})^{-1} \begin{pmatrix} I \\ B \end{pmatrix} \right],$$

we obtain

$$\text{rge} \left[ S_n \nabla \Phi(\bar{x}, \bar{y})^T S_n^T \begin{pmatrix} I \\ B^T \end{pmatrix} \right]^* = S_n \text{rge} \left[ -S_n \nabla \Phi(\bar{x}, \bar{y})^{-1} \begin{pmatrix} I \\ B \end{pmatrix} \right] = \text{rge} \left[ \nabla \Phi(\bar{x}, \bar{y})^{-1} \begin{pmatrix} I \\ B \end{pmatrix} \right].$$

As we have shown above, the latter subspace equals to  $\text{rge } Z$  with  $Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  and  $(y^*, x^*) \in (\text{rge } Z)^*$  follows.  $\square$

On the basis of Proposition 7.1 we can now establish the following characterization of strong metric regularity.

**Theorem 7.2.** Consider a mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and let  $(\bar{x}, \bar{y}) \in \text{gph } F$ .

- (i) If  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  then it is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y})$  with transformation mapping  $\Phi(x, y) = (y, x)$  and one has that  $\{\text{rge } Z \mid Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})\} \subseteq \mathcal{L}_n^{\text{reg}}$ . Further,  $F$  is SCD regular around  $(\bar{x}, \bar{y})$  and  $\text{reg } F(\bar{x}, \bar{y}) = \text{scd reg } F(\bar{x}, \bar{y})$ .

(ii) Conversely, if  $F$  is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y})$  with some transformation mapping  $\Phi$  such that  $\{\text{rge } Z \mid Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})\} \subseteq \mathcal{L}_n^{\text{reg}}$  then  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$ .

**Proof.** If  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$ , then by Theorem 2.6 it is clearly graphically Lipschitzian with the given transformation mapping  $\Phi$  and from Proposition 3.17 we obtain  $\mathcal{S}F(\bar{x}, \bar{y}) = \{\text{rge } (B, I) \mid B \in \bar{\nabla}f(\bar{y})\}$  where  $f$  denotes the Lipschitz continuous localization of  $F^{-1}$  around  $(\bar{y}, \bar{x})$ . Thus

$$\bar{\nabla}^\Phi F(\bar{x}, \bar{y}) = \left\{ \begin{pmatrix} B \\ I \end{pmatrix} \mid B \in \bar{\nabla}f(\bar{y}) \right\}$$

and consequently every matrix  $Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  is of the form  $Z = \begin{pmatrix} B \\ I \end{pmatrix}$  with  $B \in \text{conv } \bar{\nabla}f(\bar{y})$ .

From this we can easily deduce that  $L := \text{rge } Z \in \mathcal{L}_n^{\text{reg}}$  and  $B = C_L$  showing that  $F$  is SCD regular. In order to verify the formula for the modulus of strong metric regularity we use (9). Let  $\varepsilon > 0$  and consider  $(u, v) \in \text{gph } D_*F(\bar{x}, \bar{y})$  with  $\|v\| \leq 1$  and  $\|u\| \geq \text{reg } F(\bar{x}, \bar{y}) - \varepsilon$ . By Proposition 7.1 there is some  $Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  and some  $w \in \mathbb{R}^n$  with  $(u, v) = Zw$ . Thus there is  $\bar{B} \in \text{conv } \bar{\nabla}f(\bar{y})$  such that  $u = \bar{B}w$  and  $v = w$  yielding

$$\begin{aligned} \text{reg } F(\bar{x}, \bar{y}) - \varepsilon &\leq \|u\| \leq \|\bar{B}\| \|v\| \leq \|\bar{B}\| \leq \sup\{\|B\| \mid B \in \text{conv } \bar{\nabla}f(\bar{y})\} \\ &= \sup\{\|B\| \mid B \in \bar{\nabla}f(\bar{y})\} = \sup\{\|C_L\| \mid L \in \mathcal{S}F(\bar{x}, \bar{y})\} = \text{scd reg } F(\bar{x}, \bar{y}) \end{aligned}$$

by Lemma 4.5. Since  $\varepsilon > 0$  can be chosen arbitrarily small, there holds  $\text{reg } F(\bar{x}, \bar{y}) \leq \text{scd reg } F(\bar{x}, \bar{y})$ . On the other hand, we have  $\bigcup \mathcal{S}F(\bar{x}, \bar{y}) \subseteq \text{gph } D^\sharp F(\bar{x}, \bar{y}) \subseteq \text{gph } D_*F(\bar{x}, \bar{y})$  implying  $\text{scd reg } F(\bar{x}, \bar{y}) \leq \text{reg } F(\bar{x}, \bar{y})$  by Lemma 4.5 and (9). This proves (i).

The statement (ii) follows from Theorem 2.7 together with Proposition 7.1. If  $0 \in \text{gph } D_*F(\bar{x}, \bar{y})(u)$  then there is some  $Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  such that  $(u, 0) \in \text{rge } Z$  and  $u = 0$  follows from  $\text{rge } Z \in \mathcal{L}_n^{\text{reg}}$ . Similarly, if  $0 \in D^*F(\bar{x}, \bar{y})(y^*)$  then there is some  $Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  such that  $(y^*, 0) \in (\text{rge } Z)^*$ . Since  $\text{rge } Z \in \mathcal{L}_n^{\text{reg}}$ , we have  $(\text{rge } Z)^* \in \mathcal{L}_n^{\text{reg}}$  by Proposition 4.2 and  $y^* = 0$  follows. Hence, both (8) and (6) are fulfilled and strong metric regularity of  $F$  has been established.  $\square$

Consider the special case of a single-valued Lipschitzian mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  so that  $\Phi(x, y) = (x, y)$ . Then  $\bar{\nabla}^\Phi F(x, F(x)) = \{(I, B) \mid B \in \bar{\nabla}F(x)\}$  by Lemma 3.11 and therefore

$$\text{conv } \bar{\nabla}^\Phi F(x, F(x)) = \{(I, B) \mid B \in \text{conv } \bar{\nabla}F(x)\}.$$

Thus the requirement in Theorem 7.2(ii) that  $\{\text{rge } Z \mid Z \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})\} \subseteq \mathcal{L}_n^{\text{reg}}$  is equivalent to the condition that every matrix  $B$  belonging to Clarke’s generalized Jacobian  $\text{conv } \bar{\nabla}F(x)$  is nonsingular. Therefore we may consider Theorem 7.2(ii) as a generalization of Clarke’s Inverse Function Theorem, see, e.g., [6, Theorem 7.1.1], to set-valued mappings.

Note that  $\text{reg } F(\bar{x}, \bar{y}) = \text{scd reg } F(\bar{x}, \bar{y})$  whenever  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$ . This fact will not be repeated in the following results, where we present sufficient conditions for strong metric regularity.

The sufficient condition for strong metric regularity in Theorem 7.2(ii) depends on the particular basis representation  $\bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  of  $\mathcal{S}F(\bar{x}, \bar{y})$ . The next results are stated in terms of the elements  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  which do not depend on a basis.

**Corollary 7.3.** *Given  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  and a point  $(\bar{x}, \bar{y}) \in \text{gph } F$ , assume that  $\mathcal{S}F(\bar{x}, \bar{y}) = \{\bar{L}\}$  is a singleton. Then  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  if and only if  $F$  is graphically Lipschitzian of dimension  $n$  at  $(\bar{x}, \bar{y})$  and  $\bar{L} \in \mathcal{L}_n^{\text{reg}}$ .*

When  $\mathcal{S}F(\bar{x}, \bar{y})$  is a singleton and  $F$  is not graphically Lipschitzian at  $(\bar{x}, \bar{y})$  then  $F$  cannot be strongly metrically regular by Theorem 7.2(i). However, if  $F$  is SCD semismooth\* at (around)  $(\bar{x}, \bar{y})$ , then it is at least strongly metrically subregular at (around)  $(\bar{x}, \bar{y})$ . Consider the following example.

**Example 7.4.** Let  $q : \mathbb{R} \rightarrow \mathbb{R}$  be given by  $q(x) = \frac{2}{3}\text{sign}(x)|x|^{\frac{3}{2}}$ . Then  $\partial q(x) = |x|^{\frac{1}{2}}$  is not graphically Lipschitzian of dimension 1 at  $(0, 0)$  but it is an SCD mapping and SCD semismooth\*. Further,  $\mathcal{S}\partial q(0, 0) = \{\{0\} \times \mathbb{R}\}$  is a singleton and clearly  $\{0\} \times \mathbb{R} \in \mathcal{Z}_n^{\text{reg}}$ . Thus we deduce from Corollary 6.4 that  $\partial q$  is strongly metrically subregular around  $(0, 0)$ . However,  $\partial q$  is not strongly metrically regular around  $(0, 0)$  because  $\partial q^{-1}(y) = \emptyset$  for every  $y < 0$ .

We will now present a basis-independent characterization of strong metric regularity for locally maximally hypomonotone mappings.

**Theorem 7.5.** *Assume that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is locally maximally hypomonotone at  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then the following two statements are equivalent:*

- (i)  $F$  is SCD regular around  $(\bar{x}, \bar{y})$  and for every  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  the matrix  $C_L$  is positive semidefinite.
- (ii)  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  and

$$\liminf_{(x^1, y^1), (x^2, y^2) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})} \frac{\langle x^1 - x^2, y^1 - y^2 \rangle}{\|x^1 - x^2\| \|y^1 - y^2\|} \geq 0 \tag{47}$$

with the convention  $0/0 := 0$ .

**Proof.** We first prove (i) $\Rightarrow$ (ii). By Corollary 3.23 there is some  $\lambda \geq 1$  such that  $F$  is graphically Lipschitzian at  $(\bar{x}, \bar{y})$  with transformation mapping  $\Phi(x, y) = (\lambda x + y, x)$  and for every  $Z \in \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  there is a firmly nonexpansive  $n \times n$  matrix  $B$ , such that  $Z = \begin{pmatrix} B \\ I - \lambda B \end{pmatrix}$ . Since  $F$  is SCD regular around  $(\bar{x}, \bar{y})$ ,  $\pi_2(Z) = I - \lambda B$  is nonsingular and  $C_{\text{rge } Z} = B(I - \lambda B)^{-1}$ . Consider  $u \in \mathbb{R}^n$  and set  $v := (I - \lambda B)u$ . By the posed assumption,  $B(I - \lambda B)^{-1}$  is positive semidefinite and we obtain

$$0 \leq \lambda \langle v, B(I - \lambda B)^{-1}v \rangle = \lambda \langle v, Bu \rangle = \lambda \langle (I - \lambda B)u, Bu \rangle$$

implying  $\langle u, \lambda Bu \rangle \geq \|\lambda Bu\|^2$ . Thus  $\lambda B$  is firmly nonexpansive and, consequently,  $\|2\lambda B - I\| \leq 1$ . Since  $I - \lambda B$  is nonsingular, we deduce from [3, Theorem 3.3] that  $\|\lambda B\| < 1$ . Now consider  $\bar{Z} \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$ .

It follows that  $\bar{Z} = \begin{pmatrix} \bar{B} \\ I - \lambda \bar{B} \end{pmatrix}$ , where  $\bar{B}$  is some convex combination of matrices  $B_i$  with  $\|\lambda B_i\| < 1$  and

$\|2\lambda B_i - I\| \leq 1$ . It follows that  $\|\lambda \bar{B}\| < 1$  and  $\|2\lambda \bar{B} - I\| \leq 1$ . Thus  $\pi_2(\bar{Z}) = I - \lambda \bar{B}$  is nonsingular and from Proposition 4.2 we may deduce that  $\text{rge } \bar{Z} \in \mathcal{Z}_n^{\text{reg}}$ . Now strong regularity of  $F$  follows from Theorem 7.2(ii). Next we prove (47) by contraposition. Assume on the contrary that there are sequences

$(x_k^i, y_k^i) \xrightarrow{\text{gph } F} (\bar{x}, \bar{y})$ ,  $i = 1, 2$ , and some  $\eta > 0$  such that  $\langle x_k^1 - x_k^2, y_k^1 - y_k^2 \rangle < -\eta \|x_k^1 - x_k^2\| \|y_k^1 - y_k^2\|$  for all  $k$ . By local hypomonotonicity,  $\langle x_k^1 - x_k^2, y_k^1 - y_k^2 \rangle \geq -(\lambda - 1) \|x_k^1 - x_k^2\|^2$  and therefore  $\lambda > 1$  and  $\|x_k^1 - x_k^2\| \geq \frac{\eta}{\lambda - 1} \|y_k^1 - y_k^2\|$ . On the other hand, by strong metric regularity, choosing  $c > \text{reg } F(\bar{x}, \bar{y})$ , we have  $\|x_k^1 - x_k^2\| \leq c \|y_k^1 - y_k^2\|$  for all  $k$  sufficiently large. Let  $t_k := \|y_k^1 - y_k^2\| + \|x_k^1 - x_k^2\|$ . By possibly passing to a subsequence,  $(x_k^1 - x_k^2, y_k^1 - y_k^2)/t_k$  converges to some  $(u, v) \in D_*F(\bar{x}, \bar{y})$  with  $\|u\| + \|v\| = 1$ ,  $\frac{\eta}{\lambda - 1} \|v\| \leq \|u\| \leq c \|v\|$  and  $\langle u, v \rangle \leq -\eta \|u\| \|v\|$ . We deduce that both  $u$  and  $v$  are nonzero and thus  $\langle u, v \rangle < 0$ . By Proposition 7.1 there is some  $\bar{Z} \in \text{conv } \bar{\nabla}^\Phi F(\bar{x}, \bar{y})$  and some  $p \in \mathbb{R}^n$  with  $(u, v) = \bar{Z}p$ . As shown above,



$\bar{Z} = \begin{pmatrix} \bar{B} \\ I - \lambda\bar{B} \end{pmatrix}$  for some  $n \times n$  matrix  $\bar{B}$  with  $\|2\lambda\bar{B} - I\| \leq 1$ , i.e.,  $\lambda\bar{B}$  is firmly nonexpansive. It follows that

$$0 \leq \langle p, \lambda\bar{B}p \rangle - \|\lambda\bar{B}p\|^2 = \lambda\langle \bar{B}p, p - \lambda\bar{B}p \rangle = \lambda\langle u, v \rangle,$$

contradicting  $\langle u, v \rangle < 0$ . Hence the implication (i) $\Rightarrow$ (ii) is verified.

To show the reverse implication note that strong metric regularity implies SCD regularity by Theorem 7.2. We prove that  $C_L$  is positive semidefinite for every  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  by contradiction. Assume that there exists some  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  and  $p \in \mathbb{R}^n$  with  $\langle C_L p, p \rangle < 0$ . Since  $(C_L p, p) \in L \subseteq D^\sharp F(\bar{x}, \bar{y}) \subseteq D_* F(\bar{x}, \bar{y})$ , there are sequences  $t_k \downarrow 0$ ,  $(x_k, y_k) \xrightarrow{\text{gph} F} (\bar{x}, \bar{y})$  and  $(u_k, v_k) \rightarrow (C_L p, p)$  with  $(x'_k, y'_k) := (x_k, y_k) + t_k(u_k, v_k)$ . It follows that  $(x'_k - x_k)/\|x'_k - x_k\| = u_k/\|u_k\| \rightarrow C_L p$ ,  $(y'_k - y_k)/\|y'_k - y_k\| = v_k/\|v_k\| \rightarrow p/\|p\|$  and therefore

$$\lim_{k \rightarrow \infty} \frac{\langle x'_k - x_k, y'_k - y_k \rangle}{\|x'_k - x_k\| \|y'_k - y_k\|} = \frac{\langle C_L p, p \rangle}{\|C_L p\| \|p\|} < 0 \tag{48}$$

contradicting (47).  $\square$

**Corollary 7.6.** *Let  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be locally monotone around  $(\bar{x}, \bar{y}) \in \text{gph } F$ . Then the following statements are equivalent.*

- (i)  $F$  is strongly metrically regular around  $(\bar{x}, \bar{y})$ .
- (ii)  $F$  is metrically regular around  $(\bar{x}, \bar{y})$ .
- (iii)  $F$  is SCD regular around  $(\bar{x}, \bar{y})$  and locally maximally monotone at  $(\bar{x}, \bar{y})$ .

In this case, the matrices  $C_L$ ,  $L \in \mathcal{S}F(\bar{x}, \bar{y})$ , are positive semidefinite.

**Proof.** The equivalence between (i) and (ii) follows from the definitions of (strong) metric regularity and [10, Theorem 3G.5]. In view of Theorem 7.5, in order to verify (i) $\Rightarrow$ (iii), we only have to show that  $F$  is locally maximally monotone. By taking into account that  $\text{gph } F^{-1} = \{(y, x) \mid (x, y) \in \text{gph } F\}$ , it follows readily from the definition that  $F$  is locally maximally monotone at  $(\bar{x}, \bar{y})$  if and only if  $F^{-1}$  is locally maximally monotone at  $(\bar{y}, \bar{x})$ .  $F^{-1}$  has a Lipschitz continuous monotone localization and is therefore locally maximally monotone at  $(\bar{y}, \bar{x})$  by [23, Lemma 2.1]. This proves (i) $\Rightarrow$ (iii). We now claim that (iii) implies that  $C_L$  is positive semidefinite for every  $L \in \mathcal{S}F(\bar{x}, \bar{y})$ . Assuming that  $C_L$  is not positive semidefinite for some  $L \in \mathcal{S}F(\bar{x}, \bar{y})$ , the same arguments as in the proof of Theorem 7.5 can be used to obtain (48) contradicting the local monotonicity of  $F$ . Hence our claim holds true and the implication (iii) $\Rightarrow$ (i) follows from Theorem 7.5.  $\square$

**Remark 7.7.** Theorem 7.5 improves the sufficient conditions for strong metric regularity obtained by Nghia et al. [26]. E.g., in [26, Corollary 3.11] it is shown that  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is strongly metrically regular around  $(\bar{x}, \bar{y})$  if

- (a)  $F$  is locally hypomonotone at  $(\bar{x}, \bar{y})$  and
- (b)  $D^*F(\bar{x}, \bar{y})$  is positive definite in the sense that  $\langle u^*, v^* \rangle > 0$  holds for all  $u^* \in D^*F(\bar{x}, \bar{y})(v^*)$ ,  $v^* \neq 0$ .

We now show that these assumptions imply the assumptions of Theorem 7.5(i). Indeed, by the positive definiteness of the coderivative  $D^*F(\bar{x}, \bar{y})$  together with the Mordukhovich criterion (6) we may deduce that  $F$  is metrically regular around  $(\bar{x}, \bar{y})$  and therefore SCD regular as well. Further, for every  $L \in \mathcal{S}F(\bar{x}, \bar{y})$  we have  $\text{rge}(C_L^T, I) \subseteq \text{gph } D^*F(\bar{x}, \bar{y})$  by Proposition 4.2 and Remark 3.5 implying  $\langle C_L^T p, p \rangle > 0$  for all

$p$  with  $C_L^T p \neq 0$  by assumption (b). Hence,  $C_L$  is positive semidefinite. By assumption (a) there is some  $\gamma \geq 0$  such that  $\gamma I + F$  is locally monotone at  $(\bar{x}, \bar{y})$  and, since  $D^*(\gamma I + F)(\bar{x}, \gamma\bar{x} + \bar{y}) = \gamma I + D^*F(\bar{x}, \bar{y})$  is positive definite, we conclude from the Mordukhovich criterion that  $\gamma I + F$  is metrically regular around  $(\bar{x}, \gamma\bar{x} + \bar{y})$ . Thus, by Corollary 7.6, the mapping  $\gamma I + F$  is locally maximally monotone at  $(\bar{x}, \gamma\bar{x} + \bar{y})$  and  $F$  is locally maximally hypomonotone at  $(\bar{x}, \bar{y})$  by the definition. Hence, we have shown that the assumptions of Theorem 7.5(i) are weaker than those of [26, Corollary 3.11]. When we now consider, e.g., the mapping  $F(x_1, x_2) = (-x_2, x_1)$  and an arbitrary reference point  $(\bar{x}, F(\bar{x}))$ , we observe that the positive definiteness assumption (b) is not fulfilled. Nevertheless, Theorem 7.5 works well and so it in fact improves the mentioned statement in [26].

Next we turn our attention to the characterization of tilt-stable minimizers by SCD regularity of the subdifferential.

**Definition 7.8 (tilt-stable minimizers).** Let  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ , and let  $\bar{x} \in \text{dom } q$ . Then:

(i)  $\bar{x}$  is a *tilt-stable local minimizer* of  $q$  if there is a number  $\gamma > 0$  such that the mapping

$$M_\gamma(x^*) := \arg \min \{q(x) - \langle x^*, x \rangle \mid x \in \mathcal{B}_\gamma(\bar{x})\}, \quad x^* \in \mathbb{R}^n, \quad (49)$$

is single-valued and Lipschitz continuous in some neighborhood of  $\bar{x}^* = 0 \in \mathbb{R}^n$  with  $M_\gamma(0) = \{\bar{x}\}$ .

(ii) The *exact bound of tilt stability* of  $q$  at  $\bar{x}$  is defined by

$$\text{tilt}(q, \bar{x}) := \limsup_{\substack{v^*, w^* \rightarrow 0 \\ v^* \neq w^*}} \frac{\|M_\gamma(v^*) - M_\gamma(w^*)\|}{\|v^* - w^*\|}. \quad (50)$$

The theory developed in Section 5 enables us to provide a new characterization of tilt-stable local minimizers.

**Theorem 7.9.** For a function  $q : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  having  $0 \in \partial q(\bar{x})$  and such that  $q$  is both prox-regular and subdifferentially continuous at  $\bar{x}$  for  $\bar{x}^* = 0$ , the following statements are equivalent:

- (i)  $\bar{x}$  is a tilt-stable local minimizer of  $q$ .
- (ii)  $\partial q$  is SCD regular around  $(\bar{x}, 0)$  and  $C_L$  is positive semidefinite for every  $L \in \mathcal{S}\partial q(\bar{x}, 0)$ .

Further, if  $\bar{x}$  is a tilt-stable minimizer of  $q$  then  $\text{tilt}(q, \bar{x}) = \text{scd reg } \partial q(\bar{x}, 0)$ .

**Proof.** If  $\bar{x}$  is a tilt-stable minimizer, we may conclude from [28, Theorem 1.3] that the mapping  $M_\gamma$  is a single-valued Lipschitzian localization of  $\partial q^{-1}$  around  $(0, \bar{x})$  so that  $\partial q$  is strongly metrically regular around  $(\bar{x}, 0)$  and  $\text{tilt}(q, \bar{x}) = \text{reg } \partial q(\bar{x}, 0) = \text{scd reg } \partial q(\bar{x}, 0)$ . By [28, Theorem 1.3], statement (i) is equivalent to the condition

(iii) The coderivative  $D^*\partial q(\bar{x}, 0)$  is positive definite in the sense that

$$\langle v^*, u^* \rangle > 0 \quad \text{whenever} \quad u^* \in D^*F(\bar{x}, \bar{y})(v^*), \quad v^* \neq 0. \quad (51)$$

So it suffices to prove the equivalence (ii)  $\Leftrightarrow$  (iii).

*Proof that (ii)  $\Rightarrow$  (iii).* By Proposition 3.26, there is some  $\lambda > 0$  such that  $\partial q$  is graphically Lipschitzian with transformation mapping  $\Phi(x, x^*) = (x + \lambda x^*, x)$ . Further, any  $Z \in \bar{\nabla}^\Phi \partial q(\bar{x}, 0)$  satisfies  $\text{rge } Z = (\text{rge } Z)^*$

and is of the form  $Z = \begin{pmatrix} B \\ \frac{1}{\lambda}(I - B) \end{pmatrix}$ , where  $B$  is some symmetric positive semidefinite  $n \times n$  matrix. Consider  $Z \in \overline{\nabla}^\Phi \partial q(\bar{x}, 0)$  and set  $B := \pi_1(Z)$ . Since  $\partial q$  is SCD regular around  $(\bar{x}, 0)$ , the matrix  $\pi_2(Z) = \frac{1}{\lambda}(I - B)$  is nonsingular by Proposition 4.2 and  $C_{\text{rge } Z} = \lambda B(I - B)^{-1}$ . If the eigenvalues of  $B$  are denoted by  $\mu_1, \dots, \mu_n$ , the eigenvalues of  $C_L$  are  $\lambda \mu_i / (1 - \mu_i)$ ,  $i = 1, \dots, n$ , and, together with  $\mu_i \geq 0$ , we conclude that  $C_L$  is positive semidefinite if and only if  $\max \mu_i = \|B\| < 1$ . Since  $\overline{\nabla}^\Phi \partial q(\bar{x}; 0)$  is compact, it follows that  $\eta := \max\{\|\pi_1(Z)\| \mid Z \in \overline{\nabla}^\Phi \partial q(\bar{x}; 0)\} < 1$ . Now consider  $(v^*, u^*) \in \text{gph } D^*F(\bar{x}, \bar{y})$ . By Proposition 7.1 there is some  $\bar{Z} \in \text{conv } \overline{\nabla}^\Phi \partial q(\bar{x}, 0)$  such that  $(v^*, u^*) \in (\text{rge } \bar{Z})^*$ . The matrix  $\bar{B} := \pi_1(\bar{Z})$  is a convex combination of symmetric positive semidefinite matrices  $B_i$  satisfying  $\|B_i\| \leq \eta$ . Thus  $\bar{B}$  is symmetric positive semidefinite and  $\|\bar{B}\| \leq \eta < 1$ . By taking into account  $\pi_2(\bar{Z}) = \frac{1}{\lambda}(I - \bar{B})$ ,  $(\text{rge } \bar{Z})^\perp = \text{rge}(\frac{1}{\lambda}(I - \bar{B}), -\bar{B})$  and  $(\text{rge } \bar{Z})^* = \text{rge } \bar{Z}$  follows. Thus we may find some  $p \in \mathbb{R}^n$  with  $(v^*, u^*) = (\bar{B}p, \frac{1}{\lambda}(I - \bar{B})p)$  to obtain

$$\langle v^*, u^* \rangle = \frac{1}{\lambda} p^T \bar{B}(I - \bar{B})p.$$

The matrix  $\bar{B}(I - \bar{B})$  is symmetric and has eigenvalues  $\mu_i(1 - \mu_i)$ , where  $\mu_1, \dots, \mu_n$  are the eigenvalues of  $\bar{B}$ . Since  $0 \leq \mu_i \leq \eta < 1$ ,  $i = 1, \dots, n$ , the matrix  $\bar{B}(I - \bar{B})$  is positive semidefinite implying  $\langle v^*, u^* \rangle \geq 0$ . Further,  $p^T \bar{B}(I - \bar{B})p$  vanishes if and only if  $p$  is a linear combination of eigenvectors associated with the zero eigenvalues of  $\bar{B}$ , i.e.,  $\bar{B}p = v^* = 0$  and (51) follows.

*Proof that (iii)  $\Rightarrow$  (ii).* Condition (51) implies that the Mordukhovich criterion (6) is fulfilled and we may conclude that  $\partial q$  is SCD regular around  $(\bar{x}, 0)$ . Further, for every  $L \in \mathcal{S}F(\bar{x}, 0)$  we have  $L^* = \text{rge}(C_L^T, I) \subseteq \text{gph } D^* \partial q(\bar{x}, 0)$  by Proposition 4.2 and Lemma 3.7 and therefore  $(C_L^T p, p) \in \text{gph } D^* \partial q(\bar{x}, 0) \forall p$ . From (51) we deduce  $\langle C_L^T p, p \rangle > 0$  for all  $p$  with  $C_L^T p \neq 0$  and the positive semidefiniteness of  $C_L^T$  and  $C_L$  follows.  $\square$

**Example 7.10.** Consider again the mapping  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  from Example 6.5. If  $h = \nabla \phi$  for some potential  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , we see that the inclusion  $0 \in F(x)$  describes in fact the first-order optimality condition for the optimization problem

$$\min_{x \in C} \frac{1}{2} x_1^2 - \frac{1}{2} x_2^2 + \phi(x).$$

Since  $C_{TL_1}$  is not positive semidefinite, we conclude from Theorem 7.9 that  $\bar{x} = 0$  is not a tilt stable local minimizer. This is also in accordance with [28, Theorem 4.5].

### 8. Conclusion

Subspaces contained in the graph of the limiting coderivative may definitely serve as a basis for construction of suitable local approximations in the broad class of SCD multifunctions. It came to us, however, as a surprise that these subspaces and their counterparts in case of the limiting outer graphical derivative contain a lot of information about stability behavior of the considered mappings. The developed theory makes use of notions, mimicking the generalized derivatives and coderivatives in the “standard” generalized differential calculus. However, their structure is, in most cases, somewhat simpler when compared with the standard notions and so the derived new characterization of strong metric (sub)regularity and tilt stability are typically easier to work with. Finally, let us point out that the property of strong metric subregularity around the reference point, characterized via the subspaces contained in the limiting outer graphical derivative, seems to be a weak stability property ensuring the locally superlinear convergence of the semismooth\* Newton method.

## Acknowledgments

The research of the first author was supported by the Austrian Science Fund (FWF) under grant P29190-N32. The research of the second author was supported by the Grant Agency of the Czech Republic, Project 21-06569K, and the Australian Research Council, Project DP160100854. Further the authors would like to express their gratitude to an anonymous reviewer and to M. Fabián for careful reading and many important suggestions.

## References

- [1] J.P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
- [2] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
- [3] H.H. Bauschke, S.M. Moffat, X. Wang, Firmly nonexpansive mappings and maximally monotone operators: correspondence and duality, *Set-Valued Var. Anal.* 20 (2012) 131–153.
- [4] J. Bolte, A. Daniilidis, A. Lewis, Tame functions are semismooth, *Math. Program. Ser. B* 117 (2009) 5–19.
- [5] R. Cibulka, A.L. Dontchev, A.Y. Kruger, Strong metric subregularity of mappings in variational analysis and optimization, *J. Math. Anal. Appl.* 457 (2018) 1247–1282.
- [6] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.
- [7] S. Dias, G. Smirnov, On the Newton method for set-valued maps, *Nonlinear Anal.* 75 (2012) 1219–1230.
- [8] A.L. Dontchev, H. Frankowska, *On Derivative Criteria for Metric Regularity*, Computational and Analytical Mathematics, vol. 50, Springer, New York, 2013, pp. 365–374.
- [9] A.L. Dontchev, R.T. Rockafellar, Characterizations of strong regularity for variational inequalities over polyhedral convex sets, *SIAM J. Optim.* 6 (1996) 1087–1105.
- [10] A.L. Dontchev, R.T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, Heidelberg, 2014.
- [11] D. Drusvyatskiy, A.S. Lewis, Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential, *SIAM J. Optim.* 23 (2013) 256–267.
- [12] H. Gfrerer, On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs, *Set-Valued Var. Anal.* 21 (2013) 151–176.
- [13] H. Gfrerer, J.V. Outrata, On Lipschitzian properties of implicit multifunctions, *SIAM J. Optim.* 26 (2016) 2160–2189.
- [14] H. Gfrerer, J.V. Outrata, On a semismooth\* Newton method for solving generalized equations, *SIAM J. Optim.* 31 (2021) 489–517.
- [15] M.S. Gowda, Inverse and implicit function theorems for H-differentiable and semismooth functions, *Optim. Methods Softw.* 19 (2004) 443–461.
- [16] R. Henrion, J.V. Outrata, A subdifferential condition for calmness of multifunctions, *J. Math. Anal. Appl.* 258 (2001) 110–130.
- [17] T. Hoheisel, Ch. Kanzow, B.S. Mordukhovich, H. Phan, Generalized Newton’s method based on graphical derivatives, *Nonlinear Anal.* 75 (2012) 1324–1340.
- [18] N.H. Josephy, Newton’s method for generalized equations and the PIES energy model, Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin-Madison, 1979.
- [19] A. Jourani, Radiality and semismoothness, *Control Cybern.* 36 (2007) 669–680.
- [20] D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization. Regularity, Calculus, Methods and Applications*, Nonconvex Optimization and Its Applications, vol. 60, Kluwer Academic Publishers, Dordrecht, Boston, London, 2002.
- [21] A.B. Levy, Implicit multifunction theorems for the sensitivity analysis of variational conditions, *Math. Program., Ser. A* 74 (1996) 333–350.
- [22] B.S. Mordukhovich, *Variational Analysis and Applications*, Springer, Cham, 2018.
- [23] B.S. Mordukhovich, T.T.A. Nghia, Full Lipschitzian and Hölderian stability in optimization with application to mathematical programming and optimal control, *SIAM J. Optim.* 24 (2014) 1344–1381.
- [24] B.S. Mordukhovich, T.T.A. Nghia, Second-order characterizations of tilt stability with applications to nonlinear programming, *Math. Program.* 149 (2015) 83–104.
- [25] B.S. Mordukhovich, M.E. Sarabi, Generalized Newton algorithms for tilt-stable minimizers in nonsmooth optimization, *SIAM J. Optim.* 31 (2021) 1184–1214.
- [26] T.T.A. Nghia, D.T. Pham, T.T.T. Tran, On the positive definiteness of limiting coderivative for set-valued mappings, *Set-Valued Var. Anal.* 29 (2021) 283–300, <https://doi.org/10.1007/s11228-020-00547-z>.
- [27] R.A. Poliquin, R.T. Rockafellar, Prox-regular functions in variational analysis, *Trans. Am. Math. Soc.* 348 (1996) 1805–1838.
- [28] R.A. Poliquin, R.T. Rockafellar, Tilt stability of a local minimum, *SIAM J. Optim.* 8 (1998) 287–299.
- [29] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Math. Oper. Res.* 18 (1993) 227–244.
- [30] L. Qi, J. Sun, A nonsmooth version of Newton’s method, *Math. Program.* 58 (1993) 353–367.
- [31] S.M. Robinson, Strongly regular generalized equations, *Math. Oper. Res.* 5 (1980) 43–62.
- [32] S.M. Robinson, Some continuity properties of polyhedral multifunctions, *Math. Program. Stud.* 14 (1981) 206–214.
- [33] R.T. Rockafellar, Maximal monotone relations and the second derivatives of nonsmooth functions, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* 2 (1985) 167–184.

- [34] R.T. Rockafellar, R.J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
- [35] R.T. Rockafellar, Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality, submitted for publication, <https://sites.math.washington.edu/~rtr/papers/rtr256-HiddenConvexity.pdf>.