ON INNER CALMNESS*, GENERALIZED CALCULUS, AND DERIVATIVES OF THE NORMAL CONE MAPPING

Matúš Benko*

Abstract   In this paper, we study continuity and Lipschitzian properties of set-valued mappings, focusing on inner-type conditions. We introduce new notions of inner calmness* and, its relaxation, fuzzy inner calmness*. We show that polyhedral maps enjoy inner calmness* and examine (fuzzy) inner calmness* of a multiplier mapping associated with constraint systems in depth. Then we utilize these notions to develop some new rules of generalized differential calculus, mainly for the primal objects (e.g. tangent cones). In particular, we propose an exact chain rule for graphical derivatives. We apply these results to compute the derivatives of the normal cone mapping, essential e.g. for sensitivity analysis of variational inequalities.

Keywords: inner calmness*, generalized differential calculus, tangents to image sets, chain rule for graphical derivatives, normal cone mapping, polyhedral multifunctions

MSC (2010): 49J53, 49J52, 90C31

1 INTRODUCTION

The *normal cone mapping*, closely related to variational inequalities, is a standard tool of variational analysis. Among other things, it plays an important role in the formulation of optimality conditions for constrained optimization problems. Generalized derivatives of the normal cone mapping, in turn, enable us to study various stability and sensitivity properties of solutions of such optimization problems. Moreover, they also offer a pathway to optimality conditions for difficult problems, where the normal cone mapping is used to describe the constraints (mathematical programs with equilibrium constraints, bilevel programs, etc.). For more information about the normal cone mapping and related issues we refer to the standard textbooks [12, 30, 39, 44].

Recently, a lot of success has been achieved in the computation of the *graphical derivative* of the normal cone mapping $x \rightrightarrows N_\Gamma(x)$, where $\Gamma = \{x \mid g(x) \in D\}$ denotes the feasible region of a constraint system. Here, g is continuously differentiable and D is a closed convex set. For the results with polyhedral D see, e.g., [5, 7, 21, 24], while the papers [20, 26] extend the scope beyond polyhedrality. Let us mention that most of these results actually deal with more challenging cases when the feasible set depends on the parameter or even when it depends on the parameter as well as on the solution itself, see Section 5 for more details.

In general, the graphical derivative provides criteria for *isolated calmness* and the *Aubin property*, see [12, Corollary 4E.2 and Theorem 4B.2] for the precise statements and more details. On top of that, it is also used to characterize *tilt stability* of a local minimizer of a function, see [8, Theorem 3.3],

* Johannes Kepler University Linz, Institute of Computational Mathematics, 4040 Linz, Austria, and University of Vienna, Applied Mathematics and Optimization, 1090 Vienna, Austria, matus.benko@univie.ac.at

and, in the constrained case, the essential part of this criterion is precisely the graphical derivative of the normal cone mapping. Interestingly, this criterion seems to be more applicable than a similar tilt stability criterion based on the (limiting) coderivative; see e.g. [3] for the application to second-order cone programming. The reason is that the standard calculus provides an upper estimate for the limiting coderivative of the normal cone mapping, but the estimate becomes exact only under additional assumptions, see [22]. On the other hand, as suggested above, the graphical derivative can be often fully computed under very mild assumptions.

The aforementioned results regarding the computation of the graphical derivative of the normal cone mapping were derived directly from the definition. The main motivation for this paper is to better understand these computations. More precisely, we ask the question, whether there are some underlying calculus principles in play, which could be identified and brought to light.

Generalized differential calculus belongs among the most fundamental apparatus of modern variational analysis [1, 2, 30, 36, 39, 41, 44]. Arguably, more focus has been placed on the dual constructions (normal cones, subdifferentials, coderivatives) and this area is very developed and well-understood. When it comes to the primal constructions (tangent cones, subderivatives, graphical derivatives), an obvious problem appears: the lack of a suitable chain rule for graphical derivatives. A closer look shows that the problem boils down to the following issue regarding tangent cones. Given a closed set C and a continuously differentiable mapping φ, let $\bar{y} \in Q := \varphi(C)$ and consider the estimate

$$(1.1) \qquad T_Q(\bar{y}) \subset \bigcup_{(\bar{x} \in C, \varphi(\bar{x}) = \bar{y})} \nabla\varphi(\bar{x}) T_C(\bar{x}).$$

It is well-known that the opposite inclusion always holds true [44, Theorem 6.43]. The question is whether there exists a reasonable (practically applicable) assumption that guarantees this estimate. Thus, let us now comment on the standard assumptions accompanying the calculus rules.

Two main patterns can be observed throughout the most of the calculus formulas. The first one can be represented by the image rule, where, as above, the set Q under investigation is generated as the (forward) image of set C under mapping φ. The second one can be represented by the pre-image rule, where the examined set C is given as a (backward) pre-image of another given set under a given map ($C := \varphi^{-1}(Q)$). We will refer to these patterns as *forward* and *backward*, respectively.

The main assumption utilized in the backward pattern, typically called a qualification condition, is known to be the calmness of the associated perturbation mapping, often equivalently expressed via the metric subregularity of the feasibility mapping [27, 28, 31]. We point out that calmness can be viewed as an *outer* (also called upper) Lipschitzian property of set-valued mappings since, roughly speaking, it resembles outer semicontinuity with an additional Lipschitzian requirement regarding the rate of change of the map. Both calmness and metric subregularity are central notions of set-valued analysis with applications far beyond calculus. Over the years, they drew attention of several renown researchers, see the aforementioned publications as well as [11, 13, 17, 29, 42] and the references therein.

The less-developed forward pattern seems to be linked with conditions of *inner-type*. Indeed, the corresponding estimates for the dual objects are known to hold under *inner semicompactness* and *inner semicontinuity*, see [32, 38]. In order to obtain reasonable rules for the primal constructions, it turns out that one can proceed by strengthening the above continuity notions of inner-type to suitable inner (also called lower) Lipschitzian notions, such as *inner calmness*.

Interestingly, the term inner calmness was already coined in [4], where we proposed calculus rules for directional limiting normal cones and associated directional constructions. It is not very surprising, however, since directional limiting normal cone is, in a way, a primal-dual object: it consists of limits of (dual) normals along (primal) tangent directions. Moreover, inner calmness can be found in the literature also under other names, such as, e.g., Lipschitz lower semicontinuity [37] or recession with linear rate/linear recession [9, 30]. In particular, [9] contains a comprehensive study of this property.

Let us also mention the stronger notion of Lipschitz lower semicontinuity*, recently introduced in [10] and motivated by a relaxation of the Aubin property from earlier works of Klatte [34, 35].

This paper revolves around two new notions of *inner calmness** and, its relaxation, *fuzzy inner calmness**. Both these conditions are milder than inner calmness, since the latter is based on inner semicontinuity, while inner calmness* is based on (milder) inner semicompactness. All the definitions as well as some basic properties are provided in the preliminary Section 2.

The main part of the paper is Section 3, which contains a justification of the new notions. First, we show that polyhedral set-valued maps enjoy the inner calmness* property. This provides an inner counterpart to the famous result on the calmness (the upper Lipschitzness) of polyhedral maps by Robinson [42]. More importantly, in Theorem 3.9, we look into a general multiplier mapping associated with constraint systems and, under gradually strengthening assumptions, we show its inner semicompactness, fuzzy inner calmness* and inner calmness*.

In Section 4, we easily derive several calculus rules of the forward pattern. First, we prove the estimate (1.1) under fuzzy inner calmness* and then apply it to obtain a chain rule for graphical derivatives in exact form. We also provide some useful corollaries of the chain rule, such as the product rule in Theorem 4.10.

Finally, in Section 5, we return to the normal cone mapping associated with constraint systems and apply our results to compute or estimate its generalized derivatives. In particular, the computation of the graphical derivative is just a simple corollary of Theorems 3.9 and 4.10. For the most challenging setting, where the feasible region $\Gamma(p, x) = \{z \in \mid g(p, x, z) \in D\}$ depends on the parameter p as well as on the decision variable x (x typically corresponds to the solution of an underlying parametrized variational inequality), we recover the state of the art results from [5]. This suggests that our calculus based on (fuzzy) inner calmness* is applicable in quite relevant situations. Using the analogous calculus based on inner calmness is also possible if one significantly strengthens the assumptions imposed on the constraints (nevertheless, one still recovers the recent results from [24]). Using the standard result from [44, Theorem 6.43], however, seems completely out of question, since it requires convexity of the graph of the aforementioned multiplier mapping.

As a specific application, we discuss the *semismoothness** of the normal cone mapping, based on the estimates of its directional limiting coderivative. Semismoothness* is an extension of the standard semismoothness property to set-valued mappings. It was recently introduced in [25] and used to design a novel Newton method for generalized equations. The notion was then further explored e.g. in [14, 33].

The following notation is employed. Given a set $A \subset \mathbb{R}^n$, the closure and interior of A are denoted, respectively, by $\operatorname{cl} A$ and $\operatorname{int} A$, $\operatorname{sp} A$ stands for the linear hull of A and A° is the (negative) polar cone of A. A set is called cone if any nonnegative multiple of its element also belongs to the set. If $A = \{a\}$, we use just $[a]$ instead of $\operatorname{sp}\{a\}$ and $[a]^\perp$ stands for its orthogonal complement. We denote by $\operatorname{dist}(\cdot, A) := \inf_{y \in A} \|\cdot - y\|$ the usual point to set distance with the convention $\operatorname{dist}(\cdot, \emptyset) = \infty$ and $\langle \cdot, \cdot \rangle$ stands for the standard scalar product. Further, $\mathcal{B}, \mathcal{S}$ stand respectively for the closed unit ball and the unit sphere of the space in question. Given a (sufficiently) smooth function $f : \mathbb{R}^n \to \mathbb{R}$, denote its gradient and Hessian at x by $\nabla f(x)$ and $\nabla^2 f(x)$, respectively. Considering further a vector function $\varphi : \mathbb{R}^n \to \mathbb{R}^s$ with $s > 1$, denote by $\nabla \varphi(x)$ the Jacobian of φ at x, i.e., the mapping $x \to \nabla \varphi(x)$ goes from $\mathbb{R}^n$ into the space of $s \times n$ matrices, denoted by $(\mathbb{R}^s)^n$. Moreover, for a mapping $\beta : \mathbb{R}^n \to (\mathbb{R}^s)^m$ and a vector $y \in \mathbb{R}^s$, we introduce the scalarized map $\langle y, \beta \rangle : \mathbb{R}^n \to \mathbb{R}^m$ given by $\langle y, \beta \rangle(x) = \beta(x)^T y$. Following traditional patterns, given a number $\alpha \geq 0$ we denote by $o(t^\alpha)$ for $t \geq 0$ a term with the property that $o(t^\alpha)/t^\alpha \to 0$ when $t \downarrow 0$. Particularly, $o(1)$ stands for a term which converges to 0 as the variable in question goes to 0. Finally, if no confusion arises, we denote a sequence $(x_k)_{k=1}^\infty \subset \mathbb{R}^n$ simply by (x_k) and very often we also drop the brackets and write just x_k.

2 PRELIMINARIES

We begin by recalling several definitions and results from variational analysis. Let $\Omega \subset \mathbb{R}^n$ be an arbitrary closed set and $\bar{x} \in \Omega$. The *tangent* (also called *Bouligand* or *contingent*) *cone* to Ω at $\bar{x}$ is given by

$$T_\Omega(\bar{x}) := \{u \in \mathbb{R}^n \mid \exists (t_k) \downarrow 0, (u_k) \to u : \bar{x} + t_k u_k \in \Omega \,\forall k\}.$$

We denote by

$$\widehat{N}_\Omega(\bar{x}) := T_\Omega(\bar{x})^\circ = \{x^* \in \mathbb{R}^n \mid \langle x^*, u \rangle \leq 0 \,\forall u \in T_\Omega(\bar{x})\}$$

the *regular* (*Fréchet*) *normal cone* to Ω at $\bar{x}$. The *limiting* (*Mordukhovich*) *normal cone* to Ω at $\bar{x}$ is defined by

$$N_\Omega(\bar{x}) := \{x^* \in \mathbb{R}^n \mid \exists (x_k) \to \bar{x}, \ (x_k^*) \to x^* : x_k^* \in \widehat{N}_\Omega(x_k) \,\forall k\}.$$

Finally, given a direction $u \in \mathbb{R}^n$, we denote by

$$N_\Omega(\bar{x}; u) := \{x^* \in \mathbb{R}^n \mid \exists (t_k) \downarrow 0, (u_k) \to u, (x_k^*) \to x^* : x_k^* \in \widehat{N}_\Omega(\bar{x} + t_k u_k) \,\forall k\}$$

the *directional limiting normal cone* to Ω at $\bar{x}$ *in direction* u.

If $\bar{x} \notin \Omega$, we put $T_\Omega(\bar{x}) = \emptyset, \widehat{N}_\Omega(\bar{x}) = \emptyset, N_\Omega(\bar{x}) = \emptyset$ and $N_\Omega(\bar{x}; u) = \emptyset$. Further note that $N_\Omega(\bar{x}; u) = \emptyset$ whenever $u \notin T_\Omega(\bar{x})$. If Ω is convex, then $\widehat{N}_\Omega(\bar{x}) = N_\Omega(\bar{x})$ amounts to the classical normal cone in the sense of convex analysis and we will write $N_\Omega(\bar{x})$. More generally, we say that Ω is *Clarke regular* at point $\bar{x}$, provided $\widehat{N}_\Omega(\bar{x}) = N_\Omega(\bar{x})$ and we write $N_\Omega(\bar{x})$ in such case.

The following generalized derivatives of set-valued mappings are defined by means of tangents and normals to the graph of the mapping. Let $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued map with closed graph $\operatorname{gph} M := \{(x, y) \mid y \in M(x)\}$ and let $(\bar{x}, \bar{y}) \in \operatorname{gph} M$. The mapping $DM(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, defined by

$$DM(\bar{x}, \bar{y})(u) := \{v \in \mathbb{R}^m \mid (u, v) \in T_{\operatorname{gph} M}(\bar{x}, \bar{y})\},$$

is called the *graphical derivative* of M at $(\bar{x}, \bar{y})$. The mapping $\widehat{D}^* M(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$

$$\widehat{D}^* M(\bar{x}, \bar{y})(v^*) := \{u^* \in \mathbb{R}^n \mid (u^*, -v^*) \in \widehat{N}_{\operatorname{gph} M}(\bar{x}, \bar{y})\}$$

is called the *regular* (*Fréchet*) *coderivative* of M at $(\bar{x}, \bar{y})$. The mapping $D^* M(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$

$$D^* M(\bar{x}, \bar{y})(v^*) := \{u^* \in \mathbb{R}^n \mid (u^*, -v^*) \in N_{\operatorname{gph} M}(\bar{x}, \bar{y})\}$$

is called the *limiting* (*Mordukhovich*) *coderivative* of M at $(\bar{x}, \bar{y})$. Given a pair of directions $(u, v) \in \mathbb{R}^n \times \mathbb{R}^m$, the mapping $D^* M((\bar{x}, \bar{y}); (u, v)) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, given by

$$D^* M((\bar{x}, \bar{y}); (u, v))(v^*) := \{u^* \in \mathbb{R}^n \mid (u^*, -v^*) \in N_{\operatorname{gph} M}((\bar{x}, \bar{y}); (u, v))\},$$

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

In the rest of the section, we recall some well-known continuity and Lipschitzian notions for set-valued maps and introduce some new ones. As advertised in Introduction, we will mainly focus on the inner-type properties. In order to define also directional versions of these notions, we employ the following terminology. Given $x, u \in \mathbb{R}^n$, we say that a sequence x_k converges to x *from direction* u if there exist $t_k \downarrow 0$ and $u_k \to u$ with $x_k = x + t_k u_k$.

Recall that $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is *inner semicompact* at $\bar{y}$ with respect to (wrt) $\Omega \subset \mathbb{R}^m$ if for every sequence $y_k \to \bar{y}$ with $y_k \in \Omega$ there exists a subsequence K of $\mathbb{N}$ and a sequence $(x_k)_{k \in K}$ with $x_k \in S(y_k)$ for $k \in K$ converging to some $\bar{x}$. Given $\bar{x} \in S(\bar{y})$, we say that S is *inner semicontinuous* at $(\bar{y}, \bar{x})$ *wrt* Ω if for every sequence $y_k \to \bar{y}$ with $y_k \in \Omega$ there exists a sequence $x_k \to \bar{x}$ with $x_k \in S(y_k)$ for sufficiently large k. If $\Omega = \mathbb{R}^m$, we speak only about inner semicompactness at $\bar{y}$ and inner semicontinuity at

$(\bar{y}, \bar{x})$. If we restrict ourselves to sequences y_k converging to $\bar{y}$ from a fixed direction $v \in \mathbb{R}^m$, we speak of inner semicompactness and inner semicontinuity *in direction* v. For more details regarding these standard notions we refer to [38].

We point out that inner semicompactness is implied by the simpler, more intuitive, local boundedness condition [44, Definition 5.14]. In fact, local boundedness is often imposed in the development of calculus instead of inner semicompactness, see [44]. We believe, however, that the boundedness assumption is slightly misleading, since, as the name suggests, its purpose is to *restrict* the mapping, while inner semicompactness, in a sense, says the opposite - be as unbounded as you like, I just need a convergent subsequence. In Section 3, we will see the impact of these differences when dealing with a multiplier map.

For the purposes of this paper, it suffices to say that S is *outer semicontinuous* (osc) if gph S is closed, see [44, Theorem 5.7]. Hence, when dealing with derivatives of set-valued maps, given as normals or tangents to their graphs, we will use osc as a standing assumption, since we want the graphs to be closed.

In Section 4, we will be interested in tangents and normals to sets, which are actually domains of certain maps. The following lemma shows that inner semicompactness of an osc map guarantees local closedness of its domain. Clearly, local closedness, defined below, is sufficient for our needs. See [44, Theorem 5.25 (b)] for a similar result based on local boundedness.

Lemma 2.1. *Let* $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ *be osc and inner semicompact at* $\bar{y}$ *wrt* $\operatorname{dom} S := \{ y \in \mathbb{R}^m \mid S(y) \neq \emptyset \}$. *Then,* $\operatorname{dom} S$ *is locally closed around* $\bar{y}$*, i.e.,* $\operatorname{dom} S \cap V$ *is closed for some closed neighbourhood of* $\bar{y}$.

Proof. By contraposition, assume that for every k the set $\operatorname{dom} S \cap (\bar{y} + (1/k)\mathcal{B})$ is not closed, i.e., there exists $y_k \notin \operatorname{dom} S$ with $\| y_k - \bar{y} \| \leq 1/k$, together with a sequence $(y_k^l)_{l=1}^\infty \subset \operatorname{dom} S$ such that $y_k^l \to y_k$ as $l \to \infty$ and $\| y_k^l - \bar{y} \| \leq 1/k$ for all $l \in \mathbb{N}$.

Now if there exists k_0 such that for all l there exists some $x_{k_0}^l \in S(y_{k_0}^l)$ with $\| x_{k_0}^l \| \leq k_0$, then $x_{k_0}^l$ must converge to some x_{k_0} along a subsequence. Since $(y_{k_0}^l, x_{k_0}^l) \in \operatorname{gph} S$ and gph S is closed, we infer $(y_{k_0}, x_{k_0}) \in \operatorname{gph} S$, which contradicts $y_{k_0} \notin \operatorname{dom} S$. This means that for each k there exists l_k such that $x_k^{l_k} \in S(y_k^{l_k})$ implies $\| x_k^{l_k} \| > k$. Then, however, we get $\operatorname{dom} S \ni y_k^{l_k} \to \bar{y}$ as $k \to \infty$ and the inner semicompactness of S yields the existence of $x_k^{l_k} \in S(y_k^{l_k})$ converging to some $\bar{x}$ along a subsequence, contradicting $\| x_k^{l_k} \| > k$. This completes the proof. □

In [4], we needed to strengthen the above continuity properties by controlling the rate of convergence $x_k \to \bar{x}$. To this end, we came up with *inner calmness* based on inner semicontinuity. Here we also introduce the following milder concept of inner calmness* based on inner semicompactness.

Definition 2.2. A set-valued mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is called
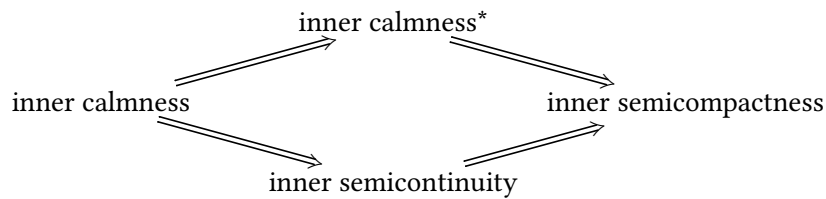
(i) *inner calm* * *at* $\bar{y} \in \mathbb{R}^m$ *wrt* $\Omega \subset \mathbb{R}^m$ if there exists $\kappa > 0$ such that for every sequence $y_k \to \bar{y}$ with $y_k \in \Omega$, there exist a subsequence K of $\mathbb{N}$, together with a sequence $(x_k)_{k \in K}$ and $\bar{x} \in \mathbb{R}^n$ with $x_k \in S(y_k)$ for $k \in K$ and

$$(2.1) \qquad\qquad\qquad \| x_k - \bar{x} \| \leq \kappa \| y_k - \bar{y} \|;$$

(ii) *inner calm at* $(\bar{y}, \bar{x}) \in \operatorname{gph} S$ *wrt* Ω if there exist $\kappa > 0$ such that for every sequence $y_k \to \bar{y}$ with $y_k \in \Omega$ there exists a sequence x_k satisfying $x_k \in S(y_k)$ and (2.1) for sufficiently large k.

As before, if we restrict ourselves to sequences y_k converging to $\bar{y}$ from a fixed direction $v \in \mathbb{R}^m$, we speak of inner calmness* and inner calmness *in direction* v.

Clearly, we have the implications:

Note that each of the four inner conditions implies that $S(y) \neq \emptyset$ for $y \in \Omega$ near $\bar{y}$. While this can be desirable in some situations, it can also be quite restrictive. For our purposes, however, we will often consider these properties wrt to the domain of S, adding no restriction at all.

We also consider one important outer Lipschitzian notion. We say that S is *calm* at $(\bar{y}, \bar{x}) \in \text{gph } S$, provided there exists $\kappa > 0$ such that for every sequence $x_k \to \bar{x}$ for which there exists a sequence y_k with $x_k \in S(y_k)$, there exists a sequence $\tilde{x}_k$ satisfying $\tilde{x}_k \in S(\bar{y})$ and

$$(2.2) \qquad \|x_k - \tilde{x}_k\| \leq \kappa \|y_k - \bar{y}\|$$

for sufficiently large k. Interestingly, in the definition of calmness, the crucial sequence is x_k in the image space $\mathbb{R}^n$. By the same token, S is called calm *in direction* $u \in \mathbb{R}^n$ if the above holds for all sequences x_k converging to $\bar{x}$ from u. This is also related to the well-known fact that calmness of S at $(\bar{y}, \bar{x})$ is equivalent to *metric subregularity* of $M := S^{-1}$ at $(\bar{x}, \bar{y})$.

Finally, we employ the following terminology. We say that S is calm at $(\bar{y}, \bar{x})$ with *constant* $\kappa \geq 0$ if κ satisfies (2.2). The infimum of all calmness constants is called the *calmness modulus* of S, which is set to be $+\infty$ if S fails to be calm. Naturally, the same applies to inner calmness* and inner calmness as well as to the directional versions of these properties.

2.1 ELEMENTARY RESULTS AND OTHER NOTIONS

Calmness and inner calmness are typically defined via neighbourhoods instead of sequences. First, we show that these definitions coincide.

Lemma 2.3. *Let $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and $(\bar{y}, \bar{x}) \in \text{gph } S$. Then*

(i) *S is inner calm at $(\bar{y}, \bar{x})$ wrt Ω with modulus $\bar{\kappa}$ if and only if $\bar{\kappa}$ is the infimum of κ over all combinations of κ and neighbourhoods V of $\bar{y}$ satisfying*

$$(2.3) \qquad \bar{x} \in S(y) + \kappa \|y - \bar{y}\|\mathcal{B} \quad \forall\, y \in V \cap \Omega;$$

(ii) *S is calm at $(\bar{y}, \bar{x})$ with modulus $\bar{\kappa}$ if and only if $\bar{\kappa}$ is the infimum of κ over all combinations of κ and neighbourhoods U of $\bar{x}$ satisfying*

$$(2.4) \qquad S(y) \cap U \subset S(\bar{y}) + \kappa \|y - \bar{y}\|\mathcal{B} \quad \forall\, y \in \mathbb{R}^m.$$

Proof. Clearly, if κ satisfies (2.3) for some neighbourhood V of $\bar{y}$, then it also satisfies (2.1) from the definition of inner calmness. On the other hand, if κ satisfies (2.1), then there exists a neighbourhood V of $\bar{y}$ such that (2.3) holds. Indeed, if not, we find a sequence $\Omega \ni y_k \to \bar{y}$ with

$$\text{dist}(\bar{x}, S(y_k)) > \kappa \|y_k - \bar{y}\|,$$

violating (2.1). Hence, the infimum defining modulus equals the infimum from the statement.

The proof for calmness is analogous.   □

Note that calmness is sometimes defined by (2.4) with $\mathbb{R}^m$ replaced by a neighbourhood of $\bar{y}$. As [12, Exercise 3H.4] clarifies, these definitions are in fact equivalent. Clearly, calmness and inner calmness are implied by the so-called Aubin property [44, Definition 9.36].

For the sake of completeness, we write down the neighbourhood-based definition of metric subregularity.

Definition 2.4. Let $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $(\bar{x}, \bar{y}) \in \operatorname{gph} M$. We say that M is *metrically subregular* at $(\bar{x}, \bar{y})$ provided there exist $\kappa > 0$ and a neighbourhood U of $\bar{x}$ such that

$$\operatorname{dist}(x, M^{-1}(\bar{y})) \le \kappa \operatorname{dist}(\bar{y}, M(x)) \quad \forall x \in U.$$

Next we look into the relation between the standard and the directional versions of the various calmness properties. We note that our interest in the directional approach stems from its successful implementation and development in recent years by Gfrerer [15].

Lemma 2.5. *Given* $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$ *and* $\Omega \subset \mathbb{R}^m$, *let* $\bar{\kappa}$ *denote the modulus of inner calmness* * *of* S *at* $\bar{y}$ *wrt* Ω *(inner calmness of* S *at* $(\bar{y}, \bar{x}) \in \operatorname{gph} S$ *wrt* Ω , *calmness of* S *at* $(\bar{y}, \bar{x})$*) and let* $\bar{\kappa}_v$ *denote the corresponding modulus in direction* $v \in \mathbb{R}^m$ *(for calmness* $\bar{\kappa}_u$ *with* $u \in \mathbb{R}^n$*). Then*

$$(2.5) \qquad \qquad \bar{\kappa} = \max_{v \in \mathcal{S}} \bar{\kappa}_v \quad (\bar{\kappa} = \max_{u \in \mathcal{S}} \bar{\kappa}_u \ \ for\ calmness).$$

Proof. Let κ by any constant of inner calmness* of S. Then, clearly, for any direction v we have $\bar{\kappa}_v \le \kappa$ and hence $\bar{\kappa}_v \le \bar{\kappa}$ by the definition of infimum and $\sup_{v \in \mathcal{S}} \bar{\kappa}_v \le \bar{\kappa}$ follows. Naturally, if $\bar{\kappa} = \infty$, the inequality holds as well.

On the other hand, let $\kappa < \bar{\kappa}$. Then there exists a sequence $y_k \in \Omega$ converging to $\bar{y}$ such that for any $\bar{x} \in \mathbb{R}^n$ there exists k_0 such that for all $x_k \in S(y_k)$ with $k \ge k_0$ one has

$$\|x_k - \bar{x}\| > \kappa \|y_k - \bar{y}\|.$$

By passing to a subsequence, however, we may assume that there exists $v \in \mathcal{S}$ such that $(y_k - \bar{y})/\|y_k - \bar{y}\| \to v$ and hence we conclude that κ can not be an inner calmness* constant of S in direction v, i.e., $\kappa \le \bar{\kappa}_v$. Thus $\bar{\kappa} \le \bar{\kappa}_v$ holds as well and (2.5) follows.

The proofs for inner calmness and calmness follow by the same steps. $\square$

We point out that (2.5) yields, in particular, that the standard (nondirectional) version of any of the calmness properties is equivalent to the validity of that property in every direction from the unit sphere. Indeed, $\bar{\kappa} = +\infty$ if and only if there exists a direction v with $\bar{\kappa}_v = +\infty$.

We will also use the following relaxed version of inner calmness*.

Definition 2.6. A set-valued mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is called *inner calm** *at* $\bar{y} \in \mathbb{R}^m$ *in direction* $v \in \mathbb{R}^m$ *wrt* $\Omega \subset \mathbb{R}^m$ *in the fuzzy sense*, if, either $v \notin T_\Omega(\bar{y})$, or there exist $\kappa_v > 0$ (*constant of fuzzy inner calmness** *in direction* v) together with sequences $y_k \in \Omega$ converging to $\bar{y}$ from v and $x_k \in S(y_k)$ and a point $\bar{x} \in \mathbb{R}^n$ such that

$$\|x_k - \bar{x}\| \le \kappa_v \|y_k - \bar{y}\|.$$

The infimum $\bar{\kappa}_v$ of all such κ_v is called the *modulus* of fuzzy inner calmness* in direction v with the convention that $\bar{\kappa}_v := 0$ if $v \notin T_\Omega(\bar{y})$. We say that S is *inner calm** *at* $\bar{y}$ *wrt* $\Omega \subset \mathbb{R}^m$ *in the fuzzy sense*, provided it is inner calm* in the fuzzy sense in every direction $v \in \mathcal{S}$.

The case $v \notin T_\Omega(\bar{y})$ is included because if there is no sequence $y_k \in \Omega$ converging to $\bar{y}$ from v, there is no requirement to meet. Moreover, it is also needed in order to make sure that inner calmness* in the fuzzy sense is implied by inner calmness*.

Note also that we do not define the (nondirectional) modulus of fuzzy inner calmness*. The reason is that, unlike in the case of the other calmness properties, it may happen that a map is inner calm* in the fuzzy sense in every direction, yet the supremum of the moduli over the unit directions blows up.
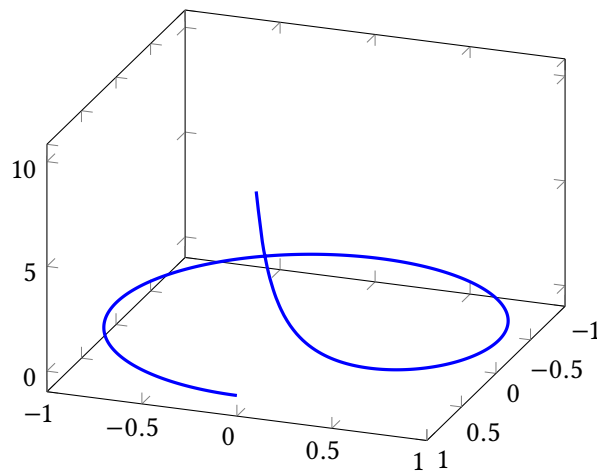
Figure 1: Sketch of set G.

Example 2.7. Consider the set $G \subset \mathbb{R}^3$ given as a curve $t \to (y_1(t), y_2(t), x(t))$ for $t \in (0, 2\pi]$ with

$$(2.6) \qquad y_1(t) = \cos(t), \ y_2(t) = \sin(t), \ x(t) = \frac{1}{t} - \frac{1}{2\pi},$$

see Figure 1. Now let $S : \mathbb{R}^2 \rightrightarrows \mathbb{R}$ be the set-valued map whose graph consists of rays starting from the origin $(0, 0, 0)$ and passing through the points of G, i.e., $\operatorname{gph} S = \{rG \mid r \geq 0\}$. Consider point $\bar{y} = (0, 0)$. It is easy to see that for any unit direction $(\cos(t), \sin(t))$ with $t \in (0, 2\pi)$, the modulus of inner calmness* and fuzzy inner calmness* coincide and equal $1/t - 1/(2\pi)$. Clearly, this quantity goes to infinity as t goes to 0, while for $t = 2\pi$ it becomes 0. In fact, this precisely entails the issue with direction $v = (1, 0)$. Indeed, on one hand it means that S is inner calm* at $\bar{y}$ in the fuzzy sense (the modulus in direction v is 0). On the other hand, from Lemma 2.5 we conclude that for the modulus of inner calmness* we have $\bar{\kappa} = \bar{\kappa}_v = +\infty$ and hence S is not inner calm* at $\bar{y}$.

Finally, in case of a single-valued mapping $\varphi : \mathbb{R}^m \to \mathbb{R}^n$, calmness at $\bar{y}$ is defined as the existence of $\kappa > 0$ and a neighbourhood V of $\bar{y}$ such that

$$(2.7) \qquad \|\varphi(y) - \varphi(\bar{y})\| \leq \kappa \|y - \bar{y}\| \quad \forall y \in V,$$

or, equivalently, the existence of $\kappa > 0$ such that (2.7), with y replaced by iterates y_k of an arbitrary sequence $y_k \to \bar{y}$, holds for sufficiently large k. If it holds for y_k converging to $\bar{y}$ from a direction v, we say φ is calm at $\bar{y}$ in direction v. Interestingly, the above definition coincides with the definition of inner calmness of φ at $(\bar{y}, \varphi(\bar{y}))$, but *not* with calmness (due to neighbourhood U) or inner calmness* (due to not requiring $\bar{x} \in S(\bar{y})$). Naturally, if we restrict ourselves to continuous mappings, all three notions coincide. Further, φ is called *Lipschitz continuous near* $\bar{y}$ if the inequality

$$\|\varphi(y) - \varphi(y')\| \leq \kappa \|y - y'\| \ \forall y, y' \in V$$

is fulfilled with $\kappa > 0$ and V being a neighbourhood of $\bar{y}$.

3 INNER SEMICOMPACTNESS AND INNER CALMNESS*

The role of inner calmness* as an assumption is highlighted in the next section dealing with the calculus rules. In this section, we discuss two interesting cases when it is satisfied. In both cases, we first prove some basic result in terms of inner semicompactness, which gets improved to inner calmness* (in the fuzzy sense) after we add suitable polyhedrality assumptions.

3.1 POLYHEDRAL SET-VALUED MAPS

We begin by the simple example showing the limitations of inner semicontinuity and inner calmness, namely, that the lack of convexity of the graph can easily lead to violation of these properties.

Example 3.1. Let $S : \mathbb{R} \rightrightarrows \mathbb{R}$ by given by

$$S(y) = \begin{cases} 0 & \text{for } y \leq 0, \\ 1 & \text{for } y \geq 0. \end{cases}$$

It is easy to see that S is not inner semicontinuous at $(\bar{y}, \bar{x}) = (0, 0)$, due to $y_k := 1/k \rightarrow \bar{y}$ and $S(y_k) = 1 \not\rightarrow 0$, or at $(\bar{y}, \bar{x}) = (0, 1)$, due to $y_k := -1/k \rightarrow \bar{y}$ and $S(y_k) = 0 \not\rightarrow 1$.

On the other hand, S is clearly inner semicompact (even inner calm*) at $\bar{y}$. Indeed, given a sequence $y_k \rightarrow \bar{y}$, we can *choose* $\bar{x}$ to be either 0 or 1, depending on which of the two values is attained by $S(y_k)$ infinitely many times.

Next, consider a map $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ whose graph is a finite union of sets G_i for $i = 1, \ldots, l$ and denote by S_i the maps with $\mathrm{gph}\, S_i = G_i$, referred to as the components of S. If G_i are closed, i.e., components S_i are osc, then so is $\mathrm{gph}\, S$ and hence S is also osc. We will now show that the properties of inner semicompactness and inner calmness* are also preserved under finite unions.

Lemma 3.2. *Given a map $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, assume that its components S_i are inner semicompact (inner calm*) at $\bar{y} \in \mathbb{R}^m$ wrt to $\mathrm{dom}\, S_i$. Then S is inner semicompact (inner calm*) at $\bar{y}$ wrt to its domain.*

Proof. Let $y_k \rightarrow \bar{y}$ with $y_k \in \mathrm{dom}\, S$. By passing to a subsequence if necessary, we may assume that there exists i with $y_k \in \mathrm{dom}\, S_i$ for all k. The inner semicompactness (inner calmness*) of S_i yields the existence of a sequence $x_k \in S_i(y_k)$ converging to some $\bar{x}$ (and $\kappa_i \geq 0$ with $\|x_k - \bar{x}\| \leq \kappa_i \|y_k - \bar{y}\|$). Since $x_k \in S(y_k)$, the claim follows (with $\kappa := \max_{i=1,\ldots,l} \kappa_i$). $\square$

The question is how to apply this lemma. It is known that mappings with convex graphs are inner semicontinuous at any $(\bar{y}, \bar{x})$ with $\bar{y}$ in the interior of the domain [44, Theorem 5.9(b)]. The following example shows, however, that even a map with closed convex graph and domain may fail to be inner semicompact wrt its domain at every point of the boundary of domain.

Example 3.3. Consider again the set $G \subset \mathbb{R}^3$ from Example 2.7. Now let $S : \mathbb{R}^2 \rightrightarrows \mathbb{R}$ be the set-valued map whose graph is the closure of the convex hull of G. Note also that the domain of S is the closed unit ball. We claim that S is not inner semicompact at $\bar{y} = (1, 0)$ wrt to $\mathrm{dom}\, S$. Indeed, consider a sequence $t_k \downarrow 0$ and set $y_k := (\cos(t_k), \sin(t_k)) \rightarrow \bar{y}$. We will show that there is no $x_k \in S(y_k)$ with $x_k < 1/t_k \rightarrow \infty$. To this end, for every k we construct a halfspace containing set G but no (y_k, x_k) with $x_k < 1/t_k$. Set $q_k := t_k^2(1/t_k - 1/(2\pi))$,

$$b_k := \left(\sin(t_k) + q_k \cos(t_k), 1 - \cos(t_k) + q_k \sin(t_k), -t_k^2(1 - \cos(t_k))\right)$$

and consider the halfspace

$$\mathcal{H}_k := \{z \in \mathbb{R}^3 \mid \langle b_k, z \rangle \leq \sin(t_k) + q_k \cos(t_k)\}.$$

Clearly, $(1, 0, 0) \in \mathcal{H}_k$. Moreover, let us explain that $G \subset \mathcal{H}_k$, i.e., that every point (y, x) of the form (2.6) also belongs to $\mathcal{H}_k$. Consider the function

$$h_k(t) := \langle b_k, (\cos(t), \sin(t), 1/t - 1/(2\pi)) \rangle.$$

A simple computation yields $h_k(t_k) = \sin(t_k) + q_k \cos(t_k)$. Thus, it suffices to show that h_k attains its global maximum over $(0, 2\pi]$ at t_k. Since

$$h'_k(t) = \langle b_k, (-\sin(t), \cos(t), -1/t^2) \rangle \quad \text{and} \quad h''_k(t) = \langle b_k, (-\cos(t), -\sin(t), 2/t^3) \rangle,$$

we get $h'_k(t_k) = 0$ and $h''_k(t_k) = -\big(\sin(t_k) + q_k + 2(1 - \cos(t_k))/t_k\big)$ and thus $h''_k(t_k) < 0$ for $t_k < \pi$.

In order to see that this maximum is not only local, note that $h_k(t) \to -\infty$ as $t \downarrow 0$ while $h_k(2\pi) = \sin(t_k) + q_k \cos(t_k) = h_k(t_k)$. Without going into details, taking into account properties of the functions defining h_k, we deduce that there is only one more stationary point of h_k in $(0, 2\pi]$, corresponding to a local minimum. Thus, the maximum at t_k is global.

Finally, since $-t_k^2(1 - \cos(t_k)) < 0$ for $t_k \in (0, 2\pi)$, a point (y_k, x) belongs to $\mathcal{H}_k$ if and only if $x \geq 1/t_k$. Since gph S is the intersection of all the closed halfspaces containing G by [43, Corollary 11.5.1], we conclude that there is no $x_k \in S(y_k)$ with $x_k < 1/t_k$. Hence, S is not inner semicompact at $\bar{y}$ wrt to dom S.

We forgo the details of applying the above lemma to a map whose graph is a union of general convex sets, since the result inevitably suffers from the limitations to the interiors of the domains of its components. Instead, we look into polyhedral mappings. Recall that a set $D \subset \mathbb{R}^s$ is said to be *convex polyhedral* if it can be represented as the intersection of finitely many halfspaces. We say that a set $E \subset \mathbb{R}^s$ is *polyhedral* if it is a union of finitely many convex polyhedral sets. A set-valued map is called (convex) polyhedral, if its graph is a (convex) polyhedral set.

In the polyhedral setting, there is no problem with the points on the boundary of the domain. Indeed, the prominent result of Walkup and Wets [45], see also [44, Example 9.35], says that a convex polyhedral map S is Lipschitz continuous on its domain, i.e., there exist $\kappa > 0$ such that

$$S(y') \subset S(y) + \kappa\|y' - y\|\mathcal{B} \quad \forall\, y, y' \in \text{dom}\, S.$$

Since this property is obviously stronger even than inner calmness at any point, we obtain the following result. For the sake of completeness, we include also the well-known result regarding (outer) calmness due to Robinson [42], who used the name *upper Lipschitzness*. We point out that the calmness below is not localized to a point $\bar{x}$, since there is no neighbourhood U as in (2.4).

Theorem 3.4 (Two-sided calmness of polyhedral maps). *Let $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a polyhedral set-valued map. Then there exists a number $\kappa > 0$ such that for every $\bar{y} \in \text{dom}\, S$, S is inner calm* with constant κ wrt $\text{dom}\, S$ at $\bar{y}$ as well as calm with constant κ at $\bar{y}$, i.e.,*

$$S(y) \subset S(\bar{y}) + \kappa\|y - \bar{y}\|\mathcal{B}$$

holds for all y near $\bar{y}$.

Note that, in inner calmness* we have found a suitable inner Lipschitzian property which, from the Lipschitzness of convex polyhedral maps, extends to polyhedral maps. The key reason is that inner calmness* is based on inner semicompactness, which is preserved under finite unions.

3.2 MULTIPLIER MAPPINGS

In [23, Proposition 4.1], Gfrerer and Outrata showed the following interesting result.

Proposition 3.5. *Let $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping having locally closed graph around $(x, \bar{y}) \in$ gph M and assume that M is metrically subregular at $(x, \bar{y})$ with modulus κ. Then*

$$(3.1) \qquad N_{M^{-1}(\bar{y})}(x) \subset \left\{ x^* \mid \exists y^* \in \kappa\|x^*\|\mathcal{B} : (x^*, y^*) \in N_{\text{gph}\, M}(x, \bar{y}) \right\}.$$

In fact, [23, Proposition 4.1] contains, apart from the additional result for the tangent cone, the stronger, directional, estimate. For our purposes, however, the important part is the *bound* $\|y^*\| \leq \kappa\|x^*\|$. Let us mention that the idea that metric subregularity yields this bound appeared already in the proof of [27, Theorem 4.1]. Moreover, similar arguments were also used in [7, Lemma 3.2] in the setting of nonlinear programs.

Corollary 3.6. *In the setting of Proposition 3.5, metric subregularity of M at $(x, \bar{y})$ implies inner semicompactness of the mapping*

$$\Lambda(x, x^*) := D^* M^{-1}(\bar{y}, x)(-x^*) = \{y^* \mid (x^*, y^*) \in N_{\mathrm{gph}\, M}(x, \bar{y})\}$$

at (x, x^*) wrt $\mathrm{gph}\, N_{M^{-1}(\bar{y})}$ *for every* $x^* \in N_{M^{-1}(\bar{y})}(x)$.

Proof. Suppose $(x_k, x_k^*) \to (x, x^*)$ with $x_k^* \in N_{M^{-1}(\bar{y})}(x_k)$. Metric subregularity of M at $(x, \bar{y})$ implies metric subregularity of M at $(x_k, \bar{y})$ for all sufficiently large k with the modulus κ' independent of k (see, e.g., [16, Lemma 2] with $u = 0$ and $\gamma = 1$). Hence, Proposition 3.5 yields the existence of $y_k^* \in \Lambda(x_k, x_k^*)$ with $\|y_k^*\| \leq \kappa' \|x_k^*\|$. In particular, by passing to a subsequence we may assume that $y_k^* \to y^*$ for some y^*, showing the inner semicompactness of Λ. $\qquad\square$

For the constraint mapping $M(x) = \varphi(x) - Q$, where $Q \subset \mathbb{R}^m$ is a closed set and $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable, (3.1) gives the standard pre-image calculus rule

$$(3.2) \qquad\qquad N_C(x) \subset \nabla\varphi(x)^T N_Q(\varphi(x))$$

where $C := M^{-1}(0) = \varphi^{-1}(Q)$, see, e.g., [4, Lemma 6.1]. Moreover, denoting $\beta(x) := \nabla\varphi(x)$, the (multiplier) mapping Λ attains the standard form

$$\Lambda(x, x^*) := \{\lambda \in N_Q(\varphi(x)) \mid \beta(x)^T \lambda = x^*\},$$

where we use λ instead of y^* to denote the multipliers.

Remark 3.7 (On inner semicompactness of the multiplier mapping). In the area of second-order analysis, one often deals with a sequence $x_k^* \in N_{\varphi^{-1}(Q)}(x_k)$. Under metric subregularity of $\varphi(x) - Q$, one gets the existence of suitable multipliers λ_k, but needs to find also a limit multiplier $\lambda \leftarrow \lambda_k$. Hence, many authors used to assume boundedness of the multipliers, for which they had to impose some stronger conditions, such as the generalized Mangasarian-Fromovitz constraint qualification (GMFCQ), see also [18] for more subtle approach. Thanks to the inner semicompactness of Λ from Corollary 3.6, however, we know that metric subregularity alone is sufficient for this task. This fact has already been utilized in several recent works, see, e.g., [3, 5, 19, 20].

We conclude this section by showing that Λ is even inner calm* (in the fuzzy sense) provided Q is polyhedral. The following lemma will be essential for our proof.

Lemma 3.8. *Let D be a convex polyhedral set and let $\bar{z}^* \in N_D(\bar{z})$. Then there exists a neighbourhood O of 0 such that*

$$(3.3) \qquad \left(\mathrm{gph}\, N_D - (\bar{z}, \bar{z}^*)\right) \cap O = \{(w, w^*) \mid w \in \mathcal{K}_D(\bar{z}, \bar{z}^*), w^* \in (\mathcal{K}_D(\bar{z}, \bar{z}^*))^\circ, \langle w, w^* \rangle = 0\} \cap O,$$

where $\mathcal{K}_D(\bar{z}, \bar{z}^*) := T_D(\bar{z}) \cap [\bar{z}^*]^\perp$ stands for the critical cone *to D at $(\bar{z}, \bar{z}^*)$.*

Proof. The *reduction lemma* [12, Lemma 2E.4] yields the existence of a neighbourhood O of 0 such that

$$(3.4) \qquad\qquad \left(\mathrm{gph}\, N_D - (\bar{z}, \bar{z}^*)\right) \cap O = \mathrm{gph}\, N_{\mathcal{K}_D(\bar{z}, \bar{z}^*)} \cap O,$$

and [12, Proposition 2A.3] gives the description of $\mathrm{gph}\, N_{\mathcal{K}_D(\bar{z}, \bar{z}^*)}$. $\qquad\square$

Theorem 3.9. *Let $(x, x^*) \in \mathrm{gph}\, N_C$ for $C = \varphi^{-1}(Q)$ with twice continuously differentiable φ and assume that the constraint mapping $M(x) = \varphi(x) - Q$ is metrically subregular at $(x, 0)$.*

   (i) If Q is Clarke regular near $\varphi(x)$ (which is the case if Q is convex), then Λ is inner semicompact at (x, x^*) wrt its domain.

(ii) *If Q is convex polyhedral, then Λ is inner calm* at (x, x^*) wrt its domain in the fuzzy sense.*

(iii) *Given $(u, u^*) \in \mathcal{S}$, if Q is convex polyhedral and the system $\varphi(\cdot) \in Q$ is non-degenerate at x in direction u, i.e.,*

$$(3.5) \qquad \beta(x)^T \lambda = 0, \ \lambda \in \mathrm{sp}\, N_{T_Q(\varphi(x))}(\nabla\varphi(x)u) \implies \lambda = 0,$$

then Λ is inner calm at (x, x^*) wrt its domain in direction (u, u^*).*

Proof. The first claim follows from Corollary 3.6 once we justify the use of dom Λ instead of gph N_C. Due to the Clarke regularity, however, we obtain $N_Q(\varphi(x')) = \widehat{N}_Q(\varphi(x'))$ for x' near x and [44, Theorem 6.14] then yields that

$$\beta(x')^T N_Q(\varphi(x')) = \beta(x')^T \widehat{N}_Q(\varphi(x')) \subset N_C(x').$$

Moreover, the opposite inclusion is also true, taking into account (3.2) and the fact that M is subregular also at $(x', 0)$ as argued in the proof of Corollary 3.6. Hence, locally around (x, x^*), one has dom $\Lambda =$ gph N_C.

To show the second one, assume that Q is convex polyhedral and consider a sequence $(x_k, x_k^*) \in$ dom Λ converging to (x, x^*) from a direction $(u, u^*) \in \mathcal{S}$, i.e.,

$$(x_k, x_k^*) = (x, x^*) + t_k(u_k, u_k^*)$$

for some $t_k \downarrow 0$ and $(u_k, u_k^*) \to (u, u^*)$. The inner semicompactness of Λ wrt dom Λ from (i) yields the existence of $\tilde{\lambda}_k \in N_Q(\varphi(x + t_k u_k))$ with

$$x^* + t_k u_k^* = \beta(x + t_k u_k)^T \tilde{\lambda}_k$$

as well as $\lambda \in N_Q(\varphi(x))$ with $x^* = \beta(x)^T \lambda$ such that $\tilde{\lambda}_k \to \lambda$. Moreover, the Taylor expansion gives

$$(3.6) \qquad \beta(x)^T \frac{\tilde{\lambda}_k - \lambda}{t_k} = u_k^* - \nabla\langle\lambda, \beta\rangle(x)u + o(1).$$

Employing Lemma 3.8 and denoting $\mathcal{K} := \mathcal{K}_Q(\varphi(x), \lambda)$ we infer

$$(3.7) \qquad \varphi(x + t_k u_k) - \varphi(x) \in \mathcal{K}, \ \tilde{\lambda}_k - \lambda \in \mathcal{K}^\circ, \ \langle\varphi(x + t_k u_k) - \varphi(x), \tilde{\lambda}_k - \lambda\rangle = 0$$

for sufficiently large k. Thus, $\mathcal{F}_k := \mathcal{K} \cap [\tilde{\lambda}_k - \lambda]^\perp$ is a *face* of the critical cone $\mathcal{K}$, see [12, p. 258], with $\varphi(x + t_k u_k) - \varphi(x) \in \mathcal{F}_k$. Since there are only finitely many faces of a polyhedral cone, we may assume that $\mathcal{F}_k \equiv \mathcal{F}$ and, in particular, we have

$$(\tilde{\lambda}_k - \lambda)/t_k \in \mathcal{K}^\circ \cap (\mathrm{sp}\,\mathcal{F})^\perp.$$

Hence, we can now invoke Hoffman's lemma [6, Theorem 2.200] to find for every k some $\eta_k \in \mathcal{K}^\circ \cap (\mathrm{sp}\,\mathcal{F})^\perp$ satisfying

$$\beta(x)^T \eta_k = \beta(x)^T \frac{\tilde{\lambda}_k - \lambda}{t_k}$$

and $\|\eta_k\| \le \alpha\|\beta(x)^T(\tilde{\lambda}_k - \lambda)/t_k\|$ for some constant $\alpha > 0$ not depending on k. Since the right-hand side of (3.6) is bounded, so is η_k and by possibly passing to a subsequence we can assume that η_k converges to some $\eta \in \mathcal{K}^\circ \cap (\mathrm{sp}\,\mathcal{F})^\perp$ satisfying

$$\beta(x)^T \eta = u^* - \nabla\langle\lambda, \beta\rangle(x)u$$

by (3.6). Since $\eta_k \in \mathcal{K}^\circ \cap (\mathrm{sp}\,\mathcal{F})^\perp$ we see that (3.7) remains true with $\tilde{\lambda}_k - \lambda$ replaced by $t_k \eta_k$. Consequently, Lemma 3.8 yields that for sufficiently large k we have

$$(\varphi(x + t_k u_k), \lambda + t_k \eta_k) = (\varphi(x), \lambda) + (\varphi(x + t_k u_k) - \varphi(x), t_k \eta_k) \in \mathrm{gph}\,N_Q.$$

Hence, setting $\lambda_k := \lambda + t_k \eta_k$, for any $\kappa_{(u,u^*)} > \|\eta\|$ we get

$$(3.8) \qquad \|\lambda_k - \lambda\| = \|\eta_k\| t_k = \frac{\|\eta_k\|}{\|(u_k, u_k^*)\|} \|(x_k, x_k^*) - (x, x^*)\| \leq \kappa_{(u,u^*)} \|(x_k, x_k^*) - (x, x^*)\|$$

for k large enough, since $\|(u_k, u_k^*)\| \to 1$ and $\|\eta_k\| \to \|\eta\|$.

Finally, denoting

$$\tilde{x}_k^* := \beta(x + t_k u_k)^T \lambda_k = \beta(x)^T \lambda + t_k(\beta(x)^T \eta_k + \nabla\langle\lambda, \beta\rangle(x)u) + o(t_k) = x^* + t_k u_k^* + o(t_k) = x_k^* + o(t_k)$$

yields, in particular, that $(x_k, \tilde{x}_k^*)$ converges to (x, x^*) from direction (u, u^*). Hence, the inner calmness* of Λ in direction (u, u^*) in the fuzzy sense with the modulus not greater than $\|\eta\|$ follows. Since direction (u, u^*) was arbitrary, Lemma 2.5 yields the inner calmness* in the fuzzy sense of Λ.

The third statement can now be shown easily. If $x_k = x$ for infinitely many k, in the previous argument we actually get

$$\beta(x)^T \eta_k = \beta(x)^T \frac{\tilde{\lambda}_k - \lambda}{t_k} = u_k^*.$$

Then, however, we obtain

$$\beta(x)^T \lambda_k = x^* + t_k u_k^* = x_k^*,$$

which means that the new multipliers λ_k correspond to the same x_k^* and this implies the inner calmness* in direction (u, u^*).

On the other hand, if $x_k \neq x$ for infinitely many k, we conduct the same arguments as in case (ii) and then show that we can in fact use the original multipliers $\tilde{\lambda}_k$. To this end, note that $\tilde{\lambda}_k - \lambda_k \in \mathrm{sp}\,N_Q(\varphi(x + t_k u_k))$ and thus [5, Proposition 2.14] says that the assumed directional non-degeneracy yields the existence of $\alpha > 0$ such that

$$\|\tilde{\lambda}_k - \lambda_k\| \leq \alpha \|\beta(x + t_k u_k)^T(\tilde{\lambda}_k - \lambda_k)\| = \alpha \|x_k^* - \tilde{x}_k^*\| = o(t_k).$$

Consequently, we obtain

$$\|\tilde{\lambda}_k - \lambda\| \leq \|\lambda_k - \lambda\| + \|\tilde{\lambda}_k - \lambda_k\| \leq \|\eta_k\| t_k + o(t_k) = (\|\eta_k\| + o(1))\, t_k$$

and we infer that $\tilde{\lambda}_k$ also satisfies the inner calmness* estimate (3.8) for any $\kappa_{(u,u^*)} > \|\eta\|$ due to $\|\eta_k\| + o(1) \to \|\eta\|$. □

Remark 3.10. Let us briefly comment on the above results.

1. We only used the fact that $\beta(x) := \nabla\varphi(x)$ in the proof of (iii). More precisely, we only used the feature of non-degeneracy from [5, Proposition 2.14]. Looking into the proof of this proposition, however, it is clear that the result remains valid under assumption (3.5) even if $\beta(x) \neq \nabla\varphi(x)$. In such case, it is not appropriate to call this assumption non-degeneracy of $\varphi(\cdot) \in Q$, naturally.

   This means that for arbitrary sufficiently smooth function $\beta(x)$ it holds that $\Lambda(x, x^*)$ is inner calm* at (x, x^*) wrt its domain in the fuzzy sense, provided Λ is inner semicompact at (x, x^*) wrt gph N_C and Q is convex polyhedral and it is even inner calm* in direction (u, u^*) if (3.5) holds as well. This enables us to handle the parametrized setting in Section 5 with ease.

2. As already mentioned, in order to replace the inner semicompactness in (i) by local boundedness, one has to impose a stronger assumption, such as GMFCQ. Moreover, in order to obtain analogous results in terms of inner semicontinuity and inner calmness, one can, e.g., assume uniqueness of the multiplier. In turn, one needs to strengthen the assumption even more and require, say, the standard non-degeneracy from [6, Formula 4.17], which corresponds to (3.5) with $u = 0$. Directional non-degeneracy was introduced in [5, Definition 2.13] and it is strictly milder than the standard one, see the example below.

3. If $(u, u^*) = (0, u^*) \in \mathcal{S}$ in (iii), (3.5) becomes the standard non-degeneracy assumption. Since $u^* \in DN_C(x, x^*)(u)$, such directions exist if and only if the normal cone mapping N_C is not isolatedly calm at (x, x^*).

4. Finally, note that our proof follows the arguments from the proof of [5, Theorem 5.3].

The following example comes from [5, Example 2.15].

Example 3.11. Consider the constraint system $\varphi(x) \leq 0$ given by

$$x_1 - x_4 \leq 0, \ -x_1 - x_4 \leq 0, \ x_2 - x_4 \leq 0, \ -x_2 - x_4 \leq 0, \ x_3 + x_1^2 - x_4 \leq 0, \ -x_3 - x_4 \leq 0$$

and the point $\bar{x} = 0$. We have $\ker \nabla \varphi(\bar{x})^T = \mathbb{R}(1, 1, 0, 0, -1, -1)^T + \mathbb{R}(0, 0, 1, 1, -1, -1)^T$. Denoting $Q := \mathbb{R}_-^6$ yields $N_Q(\varphi(\bar{x})) = \mathbb{R}_+^6$ and so sp $N_Q(\varphi(\bar{x})) = \mathbb{R}^6$, implying that the system is degenerate. Moreover, the so-called *constant rank constraint qualification* is clearly also violated at $\bar{x}$.

On the other hand, MFCQ holds $\bar{x}$, as can be seen from $\ker \nabla \varphi(\bar{x})^T \cap N_Q(\varphi(\bar{x})) = \{0\}$. More interestingly, however, let us explain that the system is actually non-degenerate in every non-zero direction. Given a direction u with $\nabla \varphi(\bar{x})u \leq 0$ and denoting $\mathcal{I}(u) := \{i \mid \nabla \varphi_i(\bar{x})u = 0\}$, we have sp $N_{T_Q(\varphi(x))}(\nabla \varphi(x)u) = \{\lambda \in \mathbb{R}^6 \mid \lambda_i = 0, i \notin \mathcal{I}(u)\}$. Thus, non-degeneracy in direction u is equivalent to the linear independence of the gradients $\nabla \varphi_i(\bar{x})$, $i \in \mathcal{I}(u)$. For $u \neq 0$, $\nabla \varphi(\bar{x})u \leq 0$ readily implies that $u_4 > 0$, since $u_4 < 0$ is not possible and $u_4 = 0$ enforces $u = 0$. Hence, $\mathcal{I}(u)$ can never contain 1 and 2 simultaneously, 3 and 4 simultaneously or 5 and 6 simultaneously, which ensures the linear independence of $\nabla \varphi_i(\bar{x})$, $i \in \mathcal{I}(u)$.

Consequently, the above theorem yields that Λ is inner calm* at $(\bar{x}, \bar{x}^*)$ wrt its domain for any $\bar{x}^* \in N_{\varphi^{-1}(Q)}(\bar{x})$.

4 SELECTED CALCULUS RULES

In this section, we show some new calculus rules based on (fuzzy) inner calmness* with focus on the primal objects (tangents and graphical derivatives).

4.1 TANGENTS AND DIRECTIONAL NORMALS TO IMAGE SETS

In this section, we deal with the rules for image sets, which provide the base for all of the remaining calculus. Given a closed set $C \subset \mathbb{R}^n$ and a continuous mapping $\varphi : \mathbb{R}^n \to \mathbb{R}^l$, set $Q := \varphi(C)$ and consider $\bar{y} \in Q$. Moreover, let $\Psi : \mathbb{R}^l \rightrightarrows \mathbb{R}^n$ be given by $\Psi(y) := \varphi^{-1}(y) \cap C$ and note that dom $\Psi = Q$ and gph $\Psi = $ gph $\varphi^{-1} \cap (\mathbb{R}^l \times C)$. Hence gph Ψ is closed by the properties of φ and C and, recalling Lemma 2.1, we suppose that Ψ is also inner semicompact at $\bar{y}$ wrt dom Ψ.

Theorem 4.1 (Tangents to image sets). *If $\varphi : \mathbb{R}^n \to \mathbb{R}^l$ is calm at some $\bar{x} \in \Psi(\bar{y})$, we have*

$$T_C(\bar{x}) \subset \{u \mid \exists v \in T_Q(\bar{y}) \text{ with } v \in D\varphi(\bar{x})u\}.$$

On the other hand, if Ψ is inner calm at $\bar{y}$ wrt dom Ψ in the fuzzy sense, then*

$$T_Q(\bar{y}) \subset \bigcup_{\bar{x} \in \Psi(\bar{y})} \{v \mid \exists u \in T_C(\bar{x}) \text{ with } v \in D\varphi(\bar{x})u\}$$

and, moreover, if there exists $\bar{x} \in \Psi(\bar{y})$ such that Ψ is inner calm at $(\bar{y}, \bar{x})$ wrt dom Ψ, *then the estimate holds with this $\bar{x}$, i.e., the union over $\Psi(\bar{y})$ is superfluous.*

Proof. Pick $u \in T_C(\bar{x})$. We will show a slightly stronger statement, namely that $D\varphi(\bar{x})u \cap T_Q(\bar{y}) \neq \emptyset$, provided φ is calm at $\bar{x}$ *in direction* u. Indeed, consider $t_k \downarrow 0$ and $u_k \to u$ with $\bar{x} + t_k u_k \in C$ and observe that

$$(4.1) \qquad\qquad Q \ni \varphi(\bar{x} + t_k u_k) = \bar{y} + t_k v_k,$$

where $v_k := (\varphi(\bar{x} + t_k u_k) - \bar{y})/t_k$ is bounded by the assumed calmness in direction u and we may assume that $v_k \to v$ for some $v \in T_Q(\bar{y})$. Moreover, (4.1) can be written as $(\bar{x} + t_k u_k, \bar{y} + t_k v_k) \in \operatorname{gph}\varphi$, showing also $v \in D\varphi(\bar{x})u$.

Consider now $v \in T_Q(\bar{y})$. The inner calmness* of Ψ at $\bar{y}$ wrt dom $\Psi = Q$ in the fuzzy sense *in direction* v yields the existence of $t_k \downarrow 0, v_k \to v$ with $\bar{y} + t_k v_k \in Q$ as well as $x_k \in C$ and $\bar{x}$ with

$$\bar{y} + t_k v_k = \varphi(x_k) \ \text{ and } \ \|x_k - \bar{x}\| \le \kappa_v t_k \|v_k\|$$

for some $\kappa_v > 0$. This means, however, that $(\bar{x} + t_k u_k, \bar{y} + t_k v_k) \in \operatorname{gph}\varphi$ for the bounded sequence $u_k := (x_k - \bar{x})/t_k$ and the existence of $u \in T_C(\bar{x})$ with $v \in D\varphi(\bar{x})u$ follows. Finally, the closedness of C and the continuity of φ imply $\bar{x} \in \Psi(\bar{y})$.

The last statement now follows easily from the definition of inner calmness. □

If φ is differentiable, $D\varphi(\bar{x})u = \{\nabla\varphi(\bar{x})u\}$ is a singleton and we obtain the following simpler estimates.

Corollary 4.2 (Tangents to image sets - differentiable case). *If φ is continuously differentiable, then*

$$T_Q(\bar{y}) \supset \bigcup_{\bar{x} \in \Psi(\bar{y})} \nabla\varphi(\bar{x}) T_C(\bar{x}).$$

The above inclusion becomes equality if Ψ is inner calm at $\bar{y}$ wrt* dom Ψ in the fuzzy sense.*

In [44, Theorem 6.43], the upper estimate for $T_Q(\bar{y})$ is obtained only in quite a special setting, namely for C convex and φ linear. As the previous section shows, our approach is applicable in a much broader context. On the other hand, our result also does not cover the estimate from [44, Theorem 6.43], which reads

$$(4.2) \qquad\qquad T_Q(\bar{y}) = \operatorname{cl}\left(\varphi(T_C(\bar{x}))\right) \text{ for every } \bar{x} \in \varphi^{-1}(\bar{y}) \cap C.$$

To see this, consider again the mapping S from Example 3.3 and set $C = \operatorname{gph} S, Q = \operatorname{dom} S$, i.e., $Q = \varphi(C)$, where φ is just the projection: $\varphi(x) = y$ for $x = (y, z) \in \mathbb{R}^2 \times \mathbb{R}$. While this setting fits into [44, Theorem 6.43], Corollary 4.2 cannot be applied. Indeed, among the standing assumptions, we ask Ψ to be inner semicompact at $\bar{y}$, but it is not since $\Psi(y) = \{y\} \times S(y)$ and S is not inner semicompact at $\bar{y}$, as clarified in Example 3.3.

Let us look more closely into the two approaches. Since it is not essential for our argument, however, we omit the details. For $\bar{y} = (1, 0)$, we clearly have $T_Q(\bar{y}) = \mathbb{R}_- \times \mathbb{R}$, since Q is the closed unit sphere. On the other hand, one can show that $S(\bar{y}) = \mathbb{R}_+$, and thus $\Psi(\bar{y}) = \varphi^{-1}(\bar{y}) \cap C = \{\bar{y}\} \times \mathbb{R}_+$, and

$$\varphi(T_C(\bar{y}, \bar{z})) = \{v \in \mathbb{R}^2 \mid \exists w \in \mathbb{R} : (v, w) \in T_C(\bar{y}, \bar{z})\} = \{(0, 0)\} \cup \{v \in \mathbb{R}^2 \mid v_1 < 0\}$$

for any $\bar{z} \in \mathbb{R}_+$. This shows the importance of the closure in (4.2). Moreover, it also shows that the estimate from Corollary 4.2 is actually false, since vectors $(0, \pm 1) \in T_Q(\bar{y})$ do not belong to $\nabla\varphi(\bar{x}) T_C(\bar{x}) = \varphi(T_C(\bar{x}))$ for any $\bar{x} = (\bar{y}, \bar{z}) \in \Psi(\bar{y})$.

Naturally, the estimates for directional normals to image sets from [4, Theorem 3.2] can also be enriched by the inner calmness* assumption. The next theorem shows how these estimates differ based on the assumption used. Given a direction $v \in \mathbb{R}^l$, let us denote

$$\Sigma(v) := \{y^* \mid D^*\varphi(\bar{x}; (u, v))(y^*) \cap N_C(\bar{x}; u) \neq \emptyset\}.$$

Trivial modifications of the proof of [4, Theorem 3.2] yield the following result.

Theorem 4.3 (Directional normals to image sets). *Consider a direction $v \in \mathbb{R}^l$ and assume that φ is Lipschitz continuous near every $\bar{x} \in \Psi(\bar{y})$. Then*

(i) if Ψ is inner semicompact at $\bar{y}$ wrt $\text{dom }\Psi$ in direction v, one has

$$N_Q(\bar{y}; v) \subset \bigcup_{\bar{x} \in \Psi(\bar{y})} \left(\bigcup_{u \in D\Psi(\bar{y}, \bar{x})(v)} \Sigma(v) \cup \bigcup_{u \in D\Psi(\bar{y}, \bar{x})(0) \cap \mathcal{S}} \Sigma(0) \right)$$

and the union over $\Psi(\bar{y})$ is superfluous if Ψ is inner semicontinuous at $(\bar{y}, \bar{x})$ wrt $\text{dom }\Psi$ in v;

(ii) if Ψ is inner calm at $\bar{y}$ wrt $\text{dom }\Psi$ in v, one has*

$$N_Q(\bar{y}; v) \subset \bigcup_{\bar{x} \in \Psi(\bar{y})} \bigcup_{u \in D\Psi(\bar{y}, \bar{x})(v)} \{y^* \mid D^*\varphi(\bar{x}; (u, v))(y^*) \cap N_C(\bar{x}; u) \neq \emptyset\}$$

and the union over $\Psi(\bar{y})$ is superfluous if Ψ is inner calm at $(\bar{y}, \bar{x})$ wrt $\text{dom }\Psi$ in v.

In fact, φ needs to be Lipschitz continuous only in relevant directions, see [4, Theorem 3.2]. The following example shows that inner calmness* in the fuzzy sense is not enough for (ii).

Example 4.4. Let $C \in \mathbb{R}^3$ be given by

$$C = \mathbb{R}_- \times \mathbb{R} \times \{0\} \cup \{x = (x_1, x_2, x_3) \mid x_1 > 0, x_2 = \sqrt{x_1}, x_3 = \sqrt{x_2}\},$$

let $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $\varphi(x) = (x_1, x_2)$ and set

$$Q := \varphi(C) = \mathbb{R}_- \times \mathbb{R} \cup \{(x_1, x_2) \mid x_1 > 0, x_2 = \sqrt{x_1}\}.$$

In order to be consistent with our notation, we use y for (x_1, x_2). First, we claim that

$$\Psi(y) = \varphi^{-1}(y) \cap C = \{x = (y, x_3) \mid x \in C\}$$

is not inner calm* at $\bar{y} = (0, 0)$ in direction $v = (0, 1)$ wrt Q, but in the fuzzy sense, it is.

Indeed, the sequence $Q \ni y^k = (1/k^2, 1/k) = (0, 0) + 1/k(1/k, 1)$ converges to $\bar{y}$ from v, but $\Psi(y^k) = (1/k^2, 1/k, 1/\sqrt{k}) \to (0, 0, 0) = \Psi(\bar{y}) = \bar{x}$ and thus

$$\|\Psi(y^k) - \Psi(\bar{y})\|/\|y^k - \bar{y}\| \geq \sqrt{k} \to \infty.$$

On the other hand, the sequence $Q \ni \tilde{y}^k = (0, 1/k) = (0, 0) + 1/k(0, 1)$ also converges to $\bar{y}$ from v with $\Psi(\tilde{y}^k) = \Psi(\bar{y})$.

We have $y^* = (-1, 0) \in N_Q(\bar{y}, v)$. The only direction $u \in T_C(\bar{x}) = \mathbb{R}_- \times \mathbb{R} \times \{0\} \cup \{0\} \times \{0\} \times \mathbb{R}_+$ with $\nabla\varphi(\bar{x})u = v$, however, is $u = (0, 1, 0)$ and $\nabla\varphi(\bar{x})^T y^* = (-1, 0, 0) \notin N_C(\bar{x}; u) = \mathbb{R}_+ \times \{0\} \times \mathbb{R}$. This shows that the estimate from (ii) does not hold.

In the remaining part of this section we will use the above results to derive other calculus rules for graphical derivatives of set-valued mappings. Let us just mention that one can also obtain the corresponding estimates, based on inner calmness*, for the dual constructions defined via directional normals, such as the directional subdifferentials of *value* (or *marginal*) functions, see [4, Theorem 4.2], or the directional coderivatives of compositions or sums of set-valued maps, see [4, Theorems 5.1 and 5.2].

4.2 CHAIN RULE AND SUM RULE FOR GRAPHICAL DERIVATIVES OF SET-VALUED MAPS

Let us begin with the chain rule. To this end, consider the mappings $S_1 : \mathbb{R}^n \rightrightarrows \mathbb{R}^m, S_2 : \mathbb{R}^m \rightrightarrows \mathbb{R}^s$ with closed graphs, i.e., S_1 and S_2 are osc, and set $S = S_2 \circ S_1$. Let $(\bar{x}, \bar{z}) \in \mathrm{gph}\, S$ and $(u, w) \in \mathbb{R}^n \times \mathbb{R}^s$ be a pair of directions. Finally, consider the "intermediate" map $\Xi : \mathbb{R}^n \times \mathbb{R}^s \rightrightarrows \mathbb{R}^m$ defined by

$$\Xi(x, z) := S_1(x) \cap S_2^{-1}(z) = \{y \in S_1(x) \mid z \in S_2(y)\}.$$

Following the approach from [44, Theorem 10.37] one has

$$(4.3) \qquad \mathrm{dom}\, \Xi = \mathrm{gph}\, S = \varphi(\mathrm{gph}\, \Xi), \quad \mathrm{gph}\, \Xi = \phi^{-1}(\mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2),$$

where $\varphi : (x, z, y) \mapsto (x, z)$ and $\phi : (x, z, y) \mapsto (x, y, y, z)$. This suggests applying first the image set rule from Corollary 4.2 and then the pre-image set rule.

Note also that

$$\Psi(x, z) := \varphi^{-1}(x, z) \cap \mathrm{gph}\, \Xi = \{(x, z, y) \mid y \in \Xi(x, z)\},$$

so it is not surprising that the fuzzy inner calmness* assumption, needed for the image set rule, can be imposed on Ξ, instead of Ψ. As before, we note that $\mathrm{gph}\, \Xi$ is closed by continuity of ϕ and osc of S_1 and S_2. Taking into account Lemma 2.1, let us assume that Ξ is inner semicompact at $(\bar{x}, \bar{z})$ wrt $\mathrm{dom}\, \Xi$.

On the other hand, for the pre-image set rule we will need metric subregularity of the mapping

$$F(x, z, y) := \mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2 - \phi(x, z, y) = \left[\begin{array}{c} \mathrm{gph}\, S_1 - (x, y) \\ \mathrm{gph}\, S_2 - (y, z) \end{array} \right].$$

Theorem 4.5 (Chain rule for graphical derivatives). *If Ξ is inner calm* at $(\bar{x}, \bar{z})$ wrt $\mathrm{dom}\, \Xi$ in the fuzzy sense, then*

$$(4.4) \qquad DS(\bar{x}, \bar{z}) \subset \bigcup_{\bar{y} \in \Xi(\bar{x}, \bar{z})} DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y}).$$

On the other hand, for any $\bar{y} \in \Xi(\bar{x}, \bar{z})$,

$$(4.5) \qquad DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y}) \subset DS(\bar{x}, \bar{z})$$

holds provided F is metrically subregular at $((\bar{x}, \bar{y}, \bar{z}), (0, 0))$ and

$$(4.6) \qquad (u, v, v, w) \in T_{\mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2}(\bar{x}, \bar{y}, \bar{y}, \bar{z}) \iff (u, v) \in T_{\mathrm{gph}\, S_1}(\bar{x}, \bar{y}),\ (v, w) \in T_{\mathrm{gph}\, S_2}(\bar{y}, \bar{z}).$$

Proof. Instead of proving the two estimates one by one, we rather propose a proof that emphasizes the role of the graphical derivative of Ξ. To this end, recall (4.3). First, we claim that, given $\bar{y} \in \Xi(\bar{x}, \bar{z})$ and $v \in D\Xi((\bar{x}, \bar{z}), \bar{y})(u, w)$, we have

$$w \in DS(\bar{x}, \bar{z})(u) \quad \text{and} \quad w \in DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y})(u).$$

Indeed, since $(u, w) = \nabla \varphi(\bar{x}, \bar{y}, \bar{z})(u, v, w)$, Corollary 4.2 yields $w \in DS(\bar{x}, \bar{z})(u)$. On the other hand, $\nabla \phi(\bar{x}, \bar{y}, \bar{z})(u, v, w) = (u, v, v, w)$ and [44, Theorem 6.31] implies $(u, v, v, w) \in T_{\mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2}(\bar{x}, \bar{y}, \bar{y}, \bar{z})$. Hence, $w \in DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y})(u)$ follows from the forward implication $\Rightarrow$ of (4.6), which always holds by [44, Proposition 6.41].

Note that we have used the image set and the pre-image set rule, but we have not needed any of the assumptions. The assumptions are in fact needed to get $\bar{y} \in \Xi(\bar{x}, \bar{z})$ and $v \in D\Xi((\bar{x}, \bar{z}), \bar{y})(u, w)$.

Consider $w \in DS(\bar{x}, \bar{z})(u)$. Thanks to the fuzzy inner calmness* of Ξ, Corollary 4.2 precisely gives $\bar{y} \in \Xi(\bar{x}, \bar{z})$ and $v \in D\Xi((\bar{x}, \bar{z}), \bar{y})(u, w)$ and (4.4) follows.

Given now $w \in DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y})(u)$, using (4.6) and then the subregularity of F (which is equivalent to the calmness of F^{-1}), [28, Proposition 1] again yields $v \in D\Xi((\bar{x}, \bar{z}), \bar{y})(u, w)$. □

We point out that the subregularity assumption and (4.6) can be replaced by asking directly for the implication

$$(4.7) \qquad w \in DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y})(u) \implies \exists v \in D\Xi((\bar{x}, \bar{z}), \bar{y})(u, w).$$

Thus, if this is satisfied for every $\bar{y} \in \Xi(\bar{x}, \bar{z})$, we get the exact chain rule

$$DS(\bar{x}, \bar{z}) = \bigcup_{\bar{y} \in \Xi(\bar{x}, \bar{z})} DS_2(\bar{y}, \bar{z}) \circ DS_1(\bar{x}, \bar{y}).$$

Note that in [44, p. 454], the authors argue that it is difficult to obtain a reasonable chain rule for graphical derivatives, since the image set and pre-image set rules for tangent cones in general work in "opposite direction", see [44, Theorems 6.31 and 6.43]. More precisely, the upper estimate of the tangent cone to image set in [44, Theorem 6.43] was obtained only under quite restrictive assumptions, see the discussion after Corollary 4.2. Here we see that it can be done using (fuzzy) inner calmness*, which can be applied in relevant situations, as the previous (and even more so the next) section shows.

Next we briefly discuss the situation when one of the mappings is single-valued.

Corollary 4.6. *If S_1 is single-valued and calm at $\bar{x}$, then $DS(\bar{x}, \bar{z}) \subset DS_2(S_1(\bar{x}), \bar{z}) \circ DS_1(\bar{x})$. If S_1 is even differentiable at $\bar{x}$, the opposite inclusion holds as well provided the map*

$$(x, z, y) \rightrightarrows (S_1(x) - y) \times (\mathrm{gph}\, S_2 - (y, z))$$

is metrically subregular at $((\bar{x}, \bar{z}, S_1(\bar{x})), (0, 0, 0))$.

Proof. If S_1 is single-valued and calm at $\bar{x}$, then Ξ, given by $\Xi(x, z) = S_1(x)$ if $z \in S_2(S_1(x))$ and $\Xi(x, z) = \emptyset$ otherwise, is single- (or empty-) valued and inner calm at $(\bar{x}, \bar{z}) \in \mathrm{gph}\, S$. This shows the first statement.

To prove the second, let us first argue that (4.6) holds for differentiable S_1. Indeed, given (u, v, w) with $v = \nabla S_1(\bar{x})(u)$ and $(v, w) \in T_{\mathrm{gph}\, S_2}(S_1(\bar{x}), \bar{z})$, consider $t_k \downarrow 0$, $(v_k, w_k) \to (v, w)$ with $\bar{z} + t_k w_k \in S_2(S_1(\bar{x}) + t_k v_k)$. We get $S_1(\bar{x} + t_k u) = S_1(\bar{x}) + t_k \tilde{v}_k$ for $\tilde{v}_k = \nabla S_1(\bar{x})u + o(1) \to v$. Thus

$$(\bar{x}, S_1(\bar{x}), S_1(\bar{x}), \bar{z}) + t_k(u, \tilde{v}_k, v_k, w_k) \in \mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2$$

and (4.6) follows. Finally, it is known and easy to check that the assumed subregularity is equivalent to the subregularity required by Theorem 4.5 (instead of $\mathrm{gph}\, S_1 - (x, y)$ we use $S_1(x) - y$). $\square$

Corollary 4.7. *Assume that S_2 is single-valued. Then (4.4) holds if Ξ is inner calm* at $(\bar{x}, \bar{z})$ wrt $\mathrm{dom}\, \Xi$ in the fuzzy sense. On the other hand, given $\bar{y} \in \Xi(\bar{x}, \bar{z})$, (4.5) holds if S_2 is differentiable at $\bar{y}$.*

Proof. We show that if S_2 is differentiable at $\bar{y}$, (4.7) holds. Consider (u, v, w) with $(u, v) \in T_{\mathrm{gph}\, S_1}(\bar{x}, \bar{y})$ and $w = \nabla S_2(\bar{y})(v)$ and let $t_k \downarrow 0$, $(u_k, v_k) \to (u, v)$ with $\bar{y} + t_k v_k \in S_1(\bar{x} + t_k u_k)$. We get $S_2(\bar{y} + t_k v_k) = S_2(\bar{y}) + t_k w_k$ for $w_k = \nabla S_2(\bar{y})v + o(1) \to w$. Thus $(\bar{x}, \bar{z}, \bar{y}) + t_k(u_k, w_k, v_k) \in \mathrm{gph}\, \Xi$ and (4.7) follows. $\square$

Before we present the sum rule, we propose an auxiliary result, interesting on its own. Consider two osc mappings $S_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^{m_i}$ for $i = 1, 2$ and let $\mathcal{P} : \mathbb{R}^n \rightrightarrows \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ be given by $\mathcal{P}(x) = S_1(x) \times S_2(x)$. Note that $\mathcal{P} = S_o \circ F_1$ for $F_1 : x \to (x, x)$ and $S_o : (q_1, q_2) \rightrightarrows S_1(q_1) \times S_2(q_2)$ and fix $(\bar{x}, \bar{y}) = (\bar{x}, (\bar{y}_1, \bar{y}_2)) \in \mathrm{gph}\, \mathcal{P}$.

Proposition 4.8. *We have*

$$D\mathcal{P}(\bar{x}, \bar{y}) \subset DS_1(\bar{x}, \bar{y}_1) \times DS_2(\bar{x}, \bar{y}_2)$$

and the opposite inclusion holds if the mapping

$$(4.8) \qquad (x, y, q) \rightrightarrows ((x, x) - (q_1, q_2)) \times (\mathrm{gph}\, S_1 - (q_1, y_1)) \times (\mathrm{gph}\, S_2 - (q_2, y_2))$$

is metrically subregular at $\big((\bar{x}, \bar{y}, (\bar{x}, \bar{x})), (0, \dots, 0)\big)$ and we have

$$(4.9) \qquad T_{\mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2}(\bar{x}, \bar{y}_1, \bar{x}, \bar{y}_2) = T_{\mathrm{gph}\, S_1}(\bar{x}, \bar{y}_1) \times T_{\mathrm{gph}\, S_2}(\bar{x}, \bar{y}_2).$$

This is true, in particular, if one of the mappings S_1, S_2 is single-valued and differentiable at $\bar{x}$.

Proof. Since $x \to (x, x)$ is differentiable and

$$\mathrm{gph}\, S_0 = \{(q_1, q_2, y_1, y_2) \mid (q_1, y_1) \in \mathrm{gph}\, S_1, (q_2, y_2) \in \mathrm{gph}\, S_2\} = \sigma(\mathrm{gph}\, S_1 \times \mathrm{gph}\, S_2),$$

where σ merely permutes the variables, Corollary 4.6 readily yields the estimates.

If one of the maps is single-valued and differentiable, (4.9) follows by the same arguments as used in the proof of Corollary 4.6. Moreover, the subregularity of mapping (4.8) is implied by its metric regularity, which reads

$$x_1^* \in D^* S_1(\bar{x}, \bar{y}_1)(0), \quad x_2^* \in D^* S_2(\bar{x}, \bar{y}_2)(0), \quad x_1^* + x_2^* = 0 \quad \Longrightarrow \quad x_1^*, x_2^* = 0,$$

see e.g. [4, page 737]. This is clearly true if one of the maps is single-valued and differentiable. $\square$

Mapping $\mathcal{P}$ naturally appears whenever one is dealing with a mapping given by some binary operation, as we will see next. In order to state the sum rule, let S_1, S_2 be as before but with $m_1 = m_2 =: m$ and consider $S = S_1 + S_2$ and $(\bar{x}, \bar{z}) \in \mathrm{gph}\, S$. Clearly S can be written via $\mathcal{P}$ as $S = F_2 \circ \mathcal{P}$ for $F_2 : (y_1, y_2) \to y_1 + y_2$. Thus, the following result follows from Proposition 4.8 and Corollary 4.7, where the corresponding "intermediate" mapping $\Xi : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^m \times \mathbb{R}^m$ is given by

$$\Xi(x, z) := \mathcal{P}(x) \cap F_2^{-1}(z) = \{y = (y_1, y_2) \in \mathbb{R}^m \times \mathbb{R}^m \mid y_1 \in S_1(x), y_2 \in S_2(x), y_1 + y_2 = z\}.$$

Theorem 4.9 (Sum rule for graphical derivatives). *If Ξ is inner calm* at $(\bar{x}, \bar{z})$ wrt $\mathrm{dom}\, \Xi$ in the fuzzy sense, then*

$$(4.10) \qquad DS(\bar{x}, \bar{z}) \subset \bigcup_{\bar{y} \in \Xi(\bar{x}, \bar{z})} DS_1(\bar{x}, \bar{y}_1) + DS_2(\bar{x}, \bar{y}_2).$$

Given $\bar{y} \in \Xi(\bar{x}, \bar{z})$ such that the mapping (4.8) is metrically subregular at $\big((\bar{x}, \bar{y}, (\bar{x}, \bar{x})), (0, \dots, 0)\big)$ and (4.9) holds, we have

$$(4.11) \qquad DS_1(\bar{x}, \bar{y}_1) + DS_2(\bar{x}, \bar{y}_2) \subset DS(\bar{x}, \bar{z}).$$

In particular, if S_1 is single-valued and differentiable, we get

$$DS(\bar{x}, \bar{z}) = \nabla S_1(\bar{x}) + DS_2(\bar{x}, \bar{z} - S_1(\bar{x})).$$

We conclude this section by another application of map $\mathcal{P}$, namely a product rule, where $S = S_1 \cdot S_2$ for an osc mapping $S_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a single-valued and differentiable $S_1 : \mathbb{R}^n \to (\mathbb{R}^m)^l$. More precisely, $S_1(x)$ is a matrix of m rows and l columns and so

$$S(x) = \bigcup_{y \in S_2(x)} S_1(x)^T y = \bigcup_{y \in S_2(x)} \langle y, S_1 \rangle(x) = F_3 \circ \mathcal{P}(x)$$

for $F_3 : (\mathbb{R}^m)^l \times \mathbb{R}^m \to \mathbb{R}^l$ given by $F(A, y) = A^T y$. Taking into account the single-valuedness of S_1, instead of $\mathcal{P}(x) \cap F_3^{-1}(z)$, we can use the following simplified "intermediate" mapping $\widetilde{\Xi} : \mathbb{R}^n \times \mathbb{R}^l \rightrightarrows \mathbb{R}^m$, given by

$$(4.12) \qquad \widetilde{\Xi}(x, z) := \{y \in S_2(x) \mid S_1(x)^T y = z\}.$$

Let $(\bar{x}, \bar{z}) \in \mathrm{gph}\, S$ and $(u, w) \in \mathbb{R}^n \times \mathbb{R}^l$ be a pair of directions. We also provide estimates for the coderivatives, since they are new and, more importantly, we will use them in the next section.

Theorem 4.10 (Product rule). *If $\widetilde{\Xi}$ is inner calm* * *at $(\bar{x}, \bar{z})$ wrt* $\operatorname{dom} \widetilde{\Xi}$ *in the fuzzy sense, then*

$$DS(\bar{x}, \bar{z})(u) = \bigcup_{\bar{y} \in \widetilde{\Xi}(\bar{x}, \bar{z})} \nabla \langle \bar{y}, S_1 \rangle(\bar{x}) u + S_1(\bar{x})^T DS_2(\bar{x}, \bar{y})(u).$$

Moreover, if $\widetilde{\Xi}$ is inner calm * *at $(\bar{x}, \bar{z})$ in direction (u, w) wrt* $\operatorname{dom} \widetilde{\Xi}$, *then*

$$D^*S((\bar{x}, \bar{z}); (u, w))(z^*) \subset \bigcup_{\bar{y} \in \widetilde{\Xi}(\bar{x}, \bar{z})} \bigcup_{v \in D\widetilde{\Xi}((\bar{x}, \bar{z}), \bar{y})(u, w)} \nabla \langle \bar{y}, S_1 \rangle(\bar{x})^T z^* + D^*S_2((\bar{x}, \bar{y}); (u, v))(S_1(\bar{x})z^*)$$

holds for all $z^* \in \mathbb{R}^l$. Finally, under just the inner semicompactness of $\widetilde{\Xi}$, we have

$$D^*S(\bar{x}, \bar{z})(z^*) \subset \bigcup_{\bar{y} \in \widetilde{\Xi}(\bar{x}, \bar{z})} \nabla \langle \bar{y}, S_1 \rangle(\bar{x})^T z^* + D^*S_2(\bar{x}, \bar{y})(S_1(\bar{x})z^*).$$

Proof. The statement about the graphical derivative again comes from Corollary 4.7 and Proposition 4.8, taking into account the simple structure of F_3.

On the basis of Theorem 4.3, mimicking the approach from [4, Theorem 5.1], one can derive the chain rule for the directional coderivatives based on inner calmness*, which looks like [4, formula (26)], but without the additional union (see the difference between the estimates in (i) and (ii) of Theorem 4.3). The proof then follows by applying [4, Corollary 5.1] and the corresponding inner calm* version of [4, Corollary 5.2]. □

5 APPLICATION: GENERALIZED DERIVATIVES OF THE NORMAL CONE MAPPING

In this section, we apply the proposed calculus rules to compute the graphical derivative of the normal cone mapping and to estimate its directional limiting coderivative. We will show that our calculus-based approach is very robust and easy to use in the case of simple constraints as well as in the parametrized setting. We then use the estimates to derive some results regarding the semismoothness* of the normal cone mapping.

5.1 SIMPLE CONSTRAINTS

Consider the simple constraint system $g(x) \in D$, where $D \subset \mathbb{R}^s$ is a convex polyhedral set and $g : \mathbb{R}^n \to \mathbb{R}^s$ is twice continuously differentiable and denote

$$(5.1) \qquad \Gamma := g^{-1}(D) = \{x \in \mathbb{R}^n \mid g(x) \in D\}.$$

Moreover, fix $\bar{x} \in \Gamma$ and assume that the constraint map $g(x) - D$ is metrically subregular at $\bar{x}$, which, in turn, means that the subregularity holds at all x near $\bar{x}$ and thus

$$N_\Gamma(x) = \nabla g(x)^T N_D(g(x)).$$

This shows that the normal cone mapping $x' \rightrightarrows N_\Gamma(x')$ can be written as the product of two maps, the single-valued ∇g and the set-valued $N_D \circ g$.

Since the set-valued part contains a composition, let us first address this issue. To this end, consider the mapping $F : \mathbb{R}^n \times \mathbb{R}^s \rightrightarrows \mathbb{R}^s \times \mathbb{R}^s$ given by

$$F(x', \lambda') := (g(x'), \lambda') - \operatorname{gph} N_D$$

and note that $\operatorname{gph}(N_D \circ g) = F^{-1}(0, 0)$.

Lemma 5.1. *Given* $\lambda \in N_D(g(x))$, *we have*

$$\eta \in D(N_D \circ g)(x, \lambda)(u) \implies \eta \in DN_D(g(x), \lambda)(\nabla g(x)u).$$

If F *is metrically subregular at* $((x, \lambda), (0, 0))$ *in direction* (u, η), *the reverse implication holds true as well and for every* $\xi^* \in \mathbb{R}^s$ *we have*

$$D^*(N_D \circ g)((x, \lambda); (u, \eta))(\xi^*) \subset \nabla g(x)^T D^* N_D((g(x), \lambda); (\nabla g(x)u, \eta))(\xi^*).$$

Proof. Since $\mathrm{gph}\,(N_D \circ g) = F^{-1}(0, 0)$, [23, Proposition 4.1] yields that

$$\eta \in D(N_D \circ g)(x, \lambda)(u) \implies (0, 0) \in DF((x, \lambda), (0, 0))(u, \eta)$$

always holds and that the reverse implication, as well as

$$D^*(N_D \circ g)((x, \lambda); (u, \eta))(\xi^*) \subset \{\xi \mid \exists (a, b) : (\xi, -\xi^*) \in D^* F(((x, \lambda), (0, 0)); ((u, \eta), 0, 0))(a, b)\},$$

hold if F is metrically subregular at $((x, \lambda), (0, 0))$ in direction (u, η). The claims now follow from the estimates of derivatives of F. Indeed, employing the sum rule from Theorem 4.9 and [4, Corollary 5.3], respectively, yields

$$(0, 0) \in DF((x, \lambda), (0, 0))(u, \eta) = (\nabla g(x)u, \eta) - T_{\mathrm{gph}\,N_D}(g(x), \lambda) \iff \eta \in DN_D(g(x), \lambda)(\nabla g(x)u)$$

and

$$(\xi, -\xi^*) \in D^* F(((x, \lambda), (0, 0)); ((u, \eta), 0, 0))(a, b)$$
$$\implies \xi \in \nabla g(x)^T D^* N_D((g(x), \lambda); (\nabla g(x)u, \eta))(\xi^*).$$

Here, we used the simple formulas for derivatives of the constant mapping $(x', \lambda') \rightrightarrows \mathrm{gph}\,N_D$ with the graph $\mathbb{R}^{n+s} \times \mathrm{gph}\,N_D$, see [4, Lemma 6.1] for more details. $\qquad\square$

Now, we can proceed with the estimates for the normal cone mapping.

Theorem 5.2. *Given* $x^* \in N_\Gamma(x)$ *for x near $\bar{x}$, for all* $u \in \mathbb{R}^n$ *we have*

$$DN_\Gamma(x, x^*)(u) \subset \bigcup_{\lambda \in \Lambda(x, x^*)} \nabla^2 \langle \lambda, g \rangle(x)u + \nabla g(x)^T N_{\mathcal{K}_D(g(x), \lambda)}(\nabla g(x)u)$$

$$= \bigcup_{\lambda \in \Lambda(x, x^*)} \nabla^2 \langle \lambda, g \rangle(x)u + N_{\mathcal{K}_\Gamma(x, x^*)}(u),$$

where

$$\Lambda(x, x^*) = \{\lambda \in N_D(g(x)) \mid \nabla g(x)^T \lambda = x^*\}.$$

The opposite inclusion holds if F *is metrically subregular at* $((x, \lambda), (0, 0))$ *for every* $\lambda \in \Lambda(x, x^*)$ *in direction* (u, η) *for every* $\eta \in N_{\mathcal{K}_D(g(x), \lambda)}(\nabla g(x)u)$. *This is true, in particular, if the following non-degeneracy condition in direction u*

$$(5.2) \qquad\qquad \nabla g(x)^T \mu = 0, \ \mu \in \mathrm{sp}\, N_{T_D(g(x))}(\nabla g(x)u) \implies \mu = 0$$

holds. Then, however, for all $w^* \in \mathbb{R}^n$ *and* $u^* \in DN_\Gamma(x, x^*)(u)$, *we also get*

$$D^* N_\Gamma((x, x^*); (u, u^*))(w^*)$$
$$\subset \bigcup_{\lambda \in \Lambda(x, x^*)} \bigcup_{\eta \in D\Lambda((x, x^*), \lambda)(u, u^*)} \nabla^2 \langle \lambda, g \rangle(x)w^* + \nabla g(x)^T D^* N_D((g(x), \lambda); (\nabla g(x)u, \eta))(\nabla g(x)w^*).$$

Proof. Theorem 3.9 (ii) yields that the multiplier mapping $\Lambda(x, x^*)$ is inner calm* at (x, x^*) wrt its domain in the fuzzy sense. Noting that Λ is precisely the "intermediate" mapping $\widetilde{\Xi}$ from (4.12) appearing in the product rule from Theorem 4.10, we obtain

$$(5.3) \qquad DN_\Gamma(x, x^*)(u) = \bigcup_{\lambda \in \Lambda(x, x^*)} \nabla^2 \langle \lambda, g \rangle(x)u + \nabla g(x)^T D(N_D \circ g)(x, \lambda)(u).$$

Next, Lemma 5.1 gives

$$(5.4) \qquad D(N_D \circ g)(x, \lambda)(u) \subset DN_D(g(x), \lambda)(\nabla g(x)u) = N_{\mathcal{K}_D(g(x), \lambda)}(\nabla g(x)u),$$

where the equality corresponds to an equivalent reformulation of the reduction lemma (3.4). Moreover, for every $\lambda \in \Lambda(x, x^*)$ it holds that

$$\mathcal{K}_\Gamma(x, x^*) = \nabla g(x)^{-1} \mathcal{K}_D(g(x), \lambda).$$

Since $u \rightrightarrows \nabla g(x)u - \mathcal{K}_D(g(x), \lambda)$ is a polyhedral map and hence metrically subregular by the Robinson's results [42, Proposition 1], we obtain

$$N_{\mathcal{K}_\Gamma(x, x^*)}(u) = \nabla g(x)^T N_{\mathcal{K}_D(g(x), \lambda)}(\nabla g(x)u)$$

and the first claim follows.

The opposite inclusion follows from the opposite inclusion in (5.4), which holds true under the subregularity assumption by Lemma 5.1.

The claim that the non-degeneracy condition implies the subregularity assumption was proven in [5, Theorem 2.11]. Moreover, from Theorem 3.9 (iii) we get that the non-degeneracy also implies the inner calmness* of Λ at (x, x^*) in direction (u, u^*) wrt its domain. Theorem 4.10 thus gives the estimate

$$D^* N_\Gamma((x, x^*); (u, u^*))(w^*)$$
$$\subset \bigcup_{\lambda \in \Lambda(x, x^*)} \bigcup_{\eta \in D\Lambda((x, x^*), \lambda)(u, u^*)} \nabla^2 \langle \lambda, g \rangle(x)w^* + D^*(N_D \circ g)((x, \lambda); (u, \eta))(\nabla g(x)w^*)$$

and using Lemma 5.1 again completes the proof. $\qquad \square$

Note that the auxiliary formula (5.3) holds without additional assumptions. Moreover, Theorem 3.9 (i) yields that the analogous estimate for the standard limiting coderivative holds for any closed set D, which is Clarke regular near $g(x)$, in particular for any closed convex set D. Finally, note also that the formula for the graphical derivative reveals that the reduction lemma for polyhedral set D, see (5.4), is carried over to set Γ with the additional curvature term $\cup_{\lambda \in \Lambda(x, x^*)} \nabla^2 \langle \lambda, g \rangle(x)u$.

As a specific application we derive the following result regarding the semismoothness* of N_Γ.

Corollary 5.3. *If the non-degeneracy condition (5.2) is satisfied for some $x, u \in \mathbb{R}^n$, then the normal cone mapping $x' \rightrightarrows N_\Gamma(x')$ is semismooth* at (x, x^*) in direction (u, u^*) for every $x^*, u^* \in \mathbb{R}^n$, i.e.,*

$$\langle u, w \rangle = \langle u^*, w^* \rangle \quad \forall (w^*, w) \in \text{gph } D^* N_\Gamma((x, x^*); (u, u^*)).$$

Proof. Clearly, if $x^* \notin N_\Gamma(x)$ or $u^* \notin DN_\Gamma(x, x^*)(u)$, then $D^* N_\Gamma((x, x^*); (u, u^*)) = \emptyset$ and there is nothing to prove. Hence, let $u^* \in DN_\Gamma(x, x^*)(u)$ and consider $w \in D^* N_\Gamma((x, x^*); (u, u^*))(w^*)$. Theorem 5.2 yields the existence of $\lambda \in \Lambda(x, x^*)$, $\eta \in DN_D(g(x), \lambda)(\nabla g(x)u)$ and

$$\zeta \in D^* N_D((g(x), \lambda); (\nabla g(x)u, \eta))(\nabla g(x)w^*)$$

such that

$$u^* = \nabla^2 \langle \lambda, g \rangle(x)u + \nabla g(x)^T \eta \text{ and } w = \nabla^2 \langle \lambda, g \rangle(x)w^* + \nabla g(x)^T \zeta.$$

Since the polyhedral map $(y, \lambda) \rightrightarrows N_D(y, \lambda)$ is semismooth* at every point of its graph, see [25, page 7], we obtain

$$
\begin{aligned}
\langle u, w \rangle &= \langle u, \nabla^2 \langle \lambda, g \rangle (x) w^* \rangle + \langle u, \nabla g(x)^T \zeta \rangle = \langle \nabla^2 \langle \lambda, g \rangle (x) u, w^* \rangle + \langle \nabla g(x) u, \zeta \rangle \\
&= \langle \nabla^2 \langle \lambda, g \rangle (x) u, w^* \rangle + \langle \eta, \nabla g(x) w^* \rangle = \langle \nabla^2 \langle \lambda, g \rangle (x) u + \nabla g(x)^T \eta, w^* \rangle = \langle u^*, w^* \rangle.
\end{aligned}
$$

$\square$

Remark 5.4. In order to show the semismoothness* of the normal come mapping, we need to show the semismoothness* in every direction (u, u^*). Clearly, the zero direction $(u, u^*) = (0, 0)$ causes no problem. In order to deal with direction $(0, u^*) \neq (0, 0)$, however, we require the non-degeneracy condition to hold for the zero direction $u = 0$, which is equivalent to the standard non-degeneracy, see Remark 3.10. This, in turn, quite significantly simplifies the situation, in particular, it implies uniqueness of λ and η, see [24]. This can be avoided if we assume that $u^* \in DN_\Gamma(x, x^*)(0)$ implies $u^* = 0$, which is equivalent with the isolated calmness of the normal cone mapping at (x, x^*). In such case we only need the non-degeneracy in all non-zero directions, which is strictly milder than the standard non-degeneracy as demonstrated in Example 3.11.

5.2 CONSTRAINTS DEPENDING ON THE PARAMETER AND THE SOLUTION

Given again a convex polyhedral set $D \subset \mathbb{R}^s$ and a twice continuously differentiable function $g : \mathbb{R}^l \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^s$, consider the feasible set

$$
(5.5) \qquad \Gamma(p, x) := \{ z \in \mathbb{R}^n \mid g(p, x, z) \in D \}
$$

depending on the parameter p as well as on the decision variable x. Here, given a reference point $(\bar{p}, \bar{x})$ with $g(\bar{p}, \bar{x}, \bar{x}) \in D$, we require the existence of $\kappa > 0$ such that for all (p, x, z) belonging to a neighbourhood of $(\bar{p}, \bar{x}, \bar{x})$ the metric inequality

$$
(5.6) \qquad \mathrm{dist}(z, \Gamma(p, x)) \leq \kappa \mathrm{dist}(g(p, x, z), D)
$$

holds. This yields, in particular, that for all $(p, x, z) \in \mathrm{gph}\,\Gamma$ sufficiently close to $(\bar{p}, \bar{x}, \bar{x})$ the mapping $g(p, x, \cdot) - D$ is metrically subregular at $(z, 0)$ with modulus κ. Hence, the normal cone mapping $(p, x) \rightrightarrows N_{\Gamma(p,x)}(x) =: \mathcal{N}^\Gamma(p, x)$ can be again written as the product

$$
\mathcal{N}^\Gamma(y) = \beta(y)^T N_D(\tilde{g}(y))
$$

for $\tilde{g}(p, x) = g(p, x, x)$, $\beta(p, x) = \nabla_3 g(p, x, x)$ and $y = (p, x)$. Given $x^* \in \mathcal{N}^\Gamma(y)$, we denote

$$
\Lambda(y, x^*) = \{ \lambda \in N_D(\tilde{g}(y)) \mid \beta(y)^T \lambda = x^* \}.
$$

Since the modulus κ of subregularity of $g(p, x, \cdot) - D$ does not depend of p and x, we infer the existence of $\lambda \in \Lambda(y, x^*)$ with $\|\lambda\| \leq \kappa \|x^*\|$ and the proof of Corollary 3.6 yields inner semicompactness of Λ at (y, x^*) wrt its domain. Moreover, Theorem 3.9 and Remark 3.10 (1.) then imply even inner calmness* of Λ in the fuzzy sense. The same arguments as in the previous case thus yield the following estimates.

Theorem 5.5. *For every $v \in \mathbb{R}^l \times \mathbb{R}^n$ we have*

$$
D\mathcal{N}^\Gamma(y, x^*)(v) \subset \bigcup_{\lambda \in \Lambda(y, x^*)} \nabla \langle \lambda, \beta \rangle (y) v + \beta(y)^T N_{\mathcal{K}_D(\tilde{g}(y), \lambda)}(\nabla \tilde{g}(y) v).
$$

On the other hand, given $v \in \mathbb{R}^l \times \mathbb{R}^n$, $\lambda \in \Lambda(y, x^*)$ and $\eta \in N_{\mathcal{K}_D(\tilde{g}(y), \lambda)}(\nabla \tilde{g}(y)(v))$, we have

$$
\nabla \langle \lambda, \beta \rangle (y) v + \beta(y)^T \eta \in D\mathcal{N}^\Gamma(y, x^*)(v),
$$

provided the mapping $(y', \lambda') \rightrightarrows (\tilde{g}(y'), \lambda') - \mathrm{gph}\, N_D$ *is metrically subregular at* $((y, \lambda), (0, 0))$ *in direction* (v, η), *in particular, if*

$$(5.7) \qquad \nabla \tilde{g}(y)^T \mu = 0, \ \mu \in \mathrm{sp}\, N_{T_D(\tilde{g}(y))}(\nabla \tilde{g}(y)v) \implies \mu = 0.$$

Moreover, if also

$$(5.8) \qquad \beta(y)^T \mu = 0, \ \mu \in \mathrm{sp}\, N_{T_D(\tilde{g}(y))}(\nabla \tilde{g}(y)v) \implies \mu = 0$$

holds, then for $u^* \in D\mathcal{N}^{\Gamma}(y, x^*)(v)$ *and arbitrary* $w^* \in \mathbb{R}^n$ *we get the estimate*

$$
\begin{aligned}
&D^* \mathcal{N}^{\Gamma}((y, x^*); (v, u^*))(w^*) \\
&\subset \bigcup_{\lambda \in \Lambda(y, x^*)} \bigcup_{\eta \in D\Lambda((y, x^*), \lambda)(v, u^*)} \nabla \langle \lambda, \beta \rangle (y)^T w^* + \nabla \tilde{g}(y)^T D^* N_D((\tilde{g}(y), \lambda); (\nabla \tilde{g}(y)v, \eta))(\beta(y)w^*).
\end{aligned}
$$

We point out that here we need to impose the two non-degeneracy conditions (5.8) and (5.7). In general, these two conditions are mutually incomparable, as can be seen from

$$\nabla \tilde{g}(p, x) = (\nabla_1 g(p, x, x), \nabla_2 g(p, x, x) + \beta(p, x)).$$

If, however, $\nabla_2 g = 0$, in particular if the feasible set Γ depends only of p but not on x, we get $\nabla \tilde{g}(p, x) = (\nabla_1 g(p, x, x), \beta(p, x))$ and, in this case, (5.8) implies (5.7). Naturally, these conditions become equivalent if also $\nabla_1 g = 0$, which covers the previous case of non-parametrized constraints.

Let us add a few comments on existing results on the topic. This model was investigated already in [40] by using the classical calculus of Mordukhovich. Recently, in [24], Gfrerer and Outrata also computed the derivatives of the normal cone mapping $(p', x') \rightrightarrows N_{\Gamma(p', x')}(x')$. Instead of (5.6) or (5.8), however, they imposed the standard non-degeneracy condition (5.8) for $v = 0$, which guarantees uniqueness of the multipliers λ and η. The reason was that they were in fact relying on inner calmness when using calculus rules, see, e.g., [24, Lemmas 4.1 and 4.2].

On the other hand, in [5], we worked with the same assumption (5.6) and proved the same result for the graphical derivative. The explicit estimation of the coderivative was, however, bypassed by addressing the related stability issues directly, see [5, Theorem 6.1].

Let us also mention the paper [20] dealing with the constraints depending on p, but not on x. This paper delivers stronger results even in a more general setting that goes beyond polyhedrality. In particular, the formula for the graphical derivative is in fact valid without the additional subregularity assumption, see [20, Theorem 3.3]. Naturally, the same applies to the non-parametrized setting.

We point out that in [5, 24] the formula for the graphical derivative is derived from [20, Theorem 5.3]. Our calculus-based approach seems to be a bit simpler and it is easily applicable regardless of whether Γ is a fixed set or it depends on p and x.

We conclude this section by the corresponding semismoothness* result.

Corollary 5.6. *Under* (5.7) *and* (5.8), *the normal cone mapping* $(p', x') \rightrightarrows N_{\Gamma(p', x')}(x')$ *is semismooth* at $(y, x^*) = (p, x, x^*)$ *in direction* $(v, u^*) = (q, u, u^*)$ *for every* $x^*, u^* \in \mathbb{R}^n$.

Proof. The proof follows by the same arguments as the proof of Corollary 5.3. We only briefly clarify that the difference between $\nabla \tilde{g}$ and β causes no problem. Indeed, looking at the above theorem and the proof of Corollary 5.3, we now get

$$\zeta \in D^* N_D((\tilde{g}(y), \lambda); (\nabla \tilde{g}(y)v, \eta))(\beta(y)w^*)$$

such that

$$u^* = \nabla \langle \lambda, \beta \rangle (y)v + \beta(y)^T \eta \ \text{ and } \ w = \nabla \langle \lambda, \beta \rangle (y)^T w^* + \nabla \tilde{g}(y)^T \zeta.$$

The computation then goes as follows

$$
\begin{aligned}
\langle v, w \rangle &= \langle v, \nabla \langle \lambda, \beta \rangle (y)^T w^* \rangle + \langle v, \nabla \tilde{g}(y)^T \zeta \rangle = \langle \nabla \langle \lambda, \beta \rangle (y) v, w^* \rangle + \langle \nabla \tilde{g}(y) v, \zeta \rangle \\
&= \langle \nabla \langle \lambda, \beta \rangle (y) v, w^* \rangle + \langle \eta, \beta(y) w^* \rangle = \langle \nabla \langle \lambda, \beta \rangle (y) v + \beta(y)^T \eta, w^* \rangle = \langle u^*, w^* \rangle.
\end{aligned}
$$

$\square$

FINAL REMARKS

Motivated by the recent success in the computation of the graphical derivative of the normal cone mapping to sets with constraint structure, in this paper, we identified the role played by (fuzzy) inner calmness* and the underlying calculus principles. We hope that it provides some helpful insights into the matter. We also hope that the new notions will turn out useful also in other circumstances. We believe that some reasonable topics for further study are readily available: Can isolated calmness be effectively used as a sufficient condition for inner calmness? Is fuzzy inner calmness* also necessary for validity of the estimate (1.1)? Can the new formulas for tangent cones yield useful estimates for regular normal cones by polarization and, in turn, enrich the theory of stationarity conditions? We plan to address these issues in the forthcoming paper.

REFERENCES

[1] J. P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, New York, 1984.

[2] J. P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 2009.

[3] M. Benko, H. Gfrerer, and B. S. Mordukhovich, Characterizations of tilt-stable minimizers in second-order cone programming, *SIAM Journal on Optimization* 29 (2019), 3100–3130, doi:10.1137/18m1213117.

[4] M. Benko, H. Gfrerer, and J. V. Outrata, Calculus for directional limiting normal cones and subdifferentials, *Set-Valued and Variational Analysis* 27 (2019), 713–745, doi:10.1007/s11228-018-0492-5.

[5] M. Benko, H. Gfrerer, and J. V. Outrata, Stability analysis for parameterized variational systems with implicit constraints, *Set-Valued and Variational Analysis* 28 (2020), 167–193, doi:10.1007/s11228-019-00516-1.

[6] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[7] N. H. Chieu and L. V. Hien, Computation of graphical derivative for a class of normal cone mappings under a very weak condition, *SIAM Journal on Optimization* 27 (2017), 190–204, doi:10.1137/16m1066816.

[8] N. H. Chieu, L. V. Hien, and T. T. A. Nghia, Characterization of tilt stability via subgradient graphical derivative with application to nonlinear programming, *SIAM Journal on Optimization* 28 (2018), 2246–2273, doi:10.1137/17m1130794.

[9] R. Cibulka, M. Fabian, and A. Y. Kruger, On semiregularity of mappings, *Journal of Mathematical Analysis and Applications* 473 (2019), 811–836, doi:10.1016/j.jmaa.2018.12.071.

[10] M. Cánovas, M. Gisbert, R. Henrion, and J. Parra, Lipschitz lower semicontinuity moduli for linear inequality systems, *Journal of Mathematical Analysis and Applications* 490 (2020), 1–21, doi:10.1016/j.jmaa.2020.124313.

[11] A. L. Dontchev and R. T. Rockafellar, Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Analysis* 12 (2004), 79–109, doi:10.1023/b:svan.0000023394.19482.30.

[12] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, Heidelberg, 2014.

[13] M. J. Fabian, R. Henrion, A. Y. Kruger, and J. V. Outrata, Error bounds: necessary and sufficient conditions, *Set-Valued and Variational Analysis* 18 (2010), 121–149, doi:10.1007/s11228-010-0133-0.

[14] M. P. Friedlander, A. Goodwin, and T. Hoheisel, From perspective maps to epigraphical projections, *arXiv* (2021), 1–31, https://arxiv.org/abs/2102.06809.

[15] H. Gfrerer, On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs, *Set-Valued and Variational Analysis* 21 (2013), 151–176, doi:10.1007/s11228-012-0220-5.

[16] H. Gfrerer, On metric pseudo-(sub)regularity of multifunctions and optimality conditions for degenerated mathematical programs, *Set-Valued and Variational Analysis* 22 (2014), 79–115, doi:10.1007/s11228-013-0266-z.

[17] H. Gfrerer and D. Klatte, Lipschitz and Hölder stability of optimization problems and generalized equations, *Mathematical Programming* 158 (2016), 35–75, doi:10.1007/s10107-015-0914-1.

[18] H. Gfrerer and B. S. Mordukhovich, Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions, *SIAM Journal on Optimization* 25 (2015), 2081–2119, doi:10.1137/15m1012608.

[19] H. Gfrerer and B. S. Mordukhovich, Robinson stability of parametric constraint systems via variational analysis, *SIAM Journal on Optimization* 27 (2017), 438–465, doi:10.1137/16m1086881.

[20] H. Gfrerer and B. S. Mordukhovich, Second-order variational analysis of parametric constraint and variational systems, *SIAM Journal on Optimization* 29 (2019), 423–453, doi:10.1137/17m1157751.

[21] H. Gfrerer and J. V. Outrata, On computation of generalized derivatives of the normal-cone mapping and their applications, *Mathematics of Operations Research* 41 (2016), 1535–1556, doi:10.1287/moor.2016.0789.

[22] H. Gfrerer and J. V. Outrata, On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications, *Optimization* 65 (2016), 671–700, doi:10.1080/02331934.2015.1066372.

[23] H. Gfrerer and J. V. Outrata, On Lipschitzian properties of implicit multifunctions, *SIAM Journal on Optimization* 26 (2016), 2160–2189, doi:10.1137/15m1052299.

[24] H. Gfrerer and J. V. Outrata, On the Aubin property of solution maps to parameterized variational systems with implicit constraints, *Optimization* 69 (2020), 1681–1701, doi:10.1080/02331934.2019.1657427.

[25] H. Gfrerer and J. V. Outrata, On a semismooth* Newton method for solving generalized equations, *SIAM Journal on Optimization* 31 (2021), 489–517, doi:10.1137/19m1257408.

[26] N. T. V. Hang, B. S. Mordukhovich, and M. E. Sarabi, Second-order variational analysis in second-order cone programming, *Mathematical Programming* 180 (2020), 75–116, doi:10.1007/s10107-018-1345-6.

[27] R. Henrion, A. Jourani, and J. V. Outrata, On the calmness of a class of multifunctions, *SIAM Journal on Optimization* 13 (2002), 603–618, doi:10.1137/s1052623401395553.

[28] R. Henrion and J. V. Outrata, Calmness of constraint systems with applications, *Mathematical Programming* 104 (2005), 437–464, doi:10.1007/s10107-005-0623-2.

[29] A. D. Ioffe, Necessary and sufficient conditions for a local minimum. 1: A reduction theorem and first order conditions, *SIAM Journal on Control and Optimization* 17 (1979), 245–250, doi:10.1137/0317019.

[30] A. D. Ioffe, *Variational Analysis of Regular Mappings*, Springer, Cham, 2017.

[31] A. D. Ioffe and J. V. Outrata, On metric and calmness qualification conditions in subdifferential calculus, *Set-Valued Analysis* 16 (2008), 199–227, doi:10.1007/s11228-008-0076-x.

[32] A. D. Ioffe and J. P. Penot, Subdifferentials of performance functions and calculus of coderivatives of set-valued mappings, *Serdica Mathematical Journal* 22 (1996), 359–384, http://hdl.handle.net/10525/612.

[33] P. D. Khanh, B. S. Mordukhovich, and V. T. Phat, A Generalized Newton Method for Subgradient Systems, *arXiv* (2020), 1–35, https://arxiv.org/abs/2009.10551.

[34] D. Klatte, A note on quantitative stability results in nonlinear programming, *Seminarbericht, Sektion Mathematik, Humboldt-Universität zu Berlin* 90 (1987), 77–86.

[35] D. Klatte, On quantitative stability for non-isolated minima, *Control and Cybernetics* 23 (1994), 183–200.

[36] D. Klatte and B. Kummer, *Nonsmooth Equations in Optimization*, Kluwer Academic, Dordrecht, 2002.

[37] D. Klatte and B. Kummer, On calmness of the argmin mapping in parametric optimization problems, *Journal of Optimization Theory and Applications* 165 (2015), 708–719, doi:10.1007/s10957-014-0643-2.

[38] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, Vol. I: Basic Theory, Vol. II: Applications*, Springer, Berlin, 2006.

[39] B. S. Mordukhovich, *Variational Analysis and Applications*, Springer, Cham, 2018.

[40] B. S. Mordukhovich and J. V. Outrata, Coderivative analysis of quasi-variational inequalities with applications to stability and optimization, *SIAM Journal on Optimization* 18 (2007), 389–412, doi:10.1137/060665609.

[41] J. P. Penot, *Calculus Without Derivatives*, Springer, New York, 2013.

[42] S. M. Robinson, Some continuity properties of polyhedral multifunctions, in *Mathematical Programming at Oberwolfach*, H. König, B. Korte, and K. Ritter (eds.), Springer, Berlin, 1981, 206–214.

[43] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[44] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[45] D. W. Walkup and R. J. B. Wets, A Lipschitz characterization of convex polyhedra, *Proceedings of the American Mathematical Society* 23 (1969), 167–173, doi:10.2307/2037511.