Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

On the co-derivative of normal cone mappings to inequality systems

René Henrion[1], Jiří Outrata[2], Thomas Surowiec[3]

submitted: April 28, 2008

[1] Weierstrass Institute
for Applied Analysis
and Stochastics
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: henrion@wias-berlin.de

[2] Institute of
Information Theory and
Automation Theory
18208 Praha 8
Czech Republic
E-Mail: outrata@utia.cas.cz

[3] Humboldt University Berlin
Institute of Mathematics
Unter den Linden 6
10099 Berlin
Germany
E-Mail: surowiec@math.hu-berlin.de

No. 1321
Berlin 2008

2000 *Mathematics Subject Classification.* 90C30, 49J53.

Key words and phrases. Mordukhovich co-derivative, normal cone mapping, calmness.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

The paper deals with co-derivative formulae for normal cone mappings to smooth inequality systems. Both, the regular (Linear Independence Constraint Qualification satisfied) and nonregular (Mangasarian-Fromovitz Constraint Qualification satisfied) case are considered. A major part of the results relies on general transformation formulae previously obtained by Mordukhovich and Outrata. This allows to derive exact formulae for general smooth, regular and polyhedral, possibly nonregular systems. In the nonregular, nonpolyhedral case a generalized transformation formula by Mordukhovich and Outrata applies, however a major difficulty consists in checking a calmness condition of a certain multivalued mapping. The paper provides a translation of this condition in terms of much easier to verify constraint qualifications. A series of examples illustrates the use and comparison of the presented formulae.

1 Introduction

The Mordukhovich co-derivative has become an important tool for the characterization of stability and optimality in variational analysis. We refer to the basic monograph [7] for definitions, properties, calculus rules and applications of this object. When dealing with generalized equations or variational inequalities, the multivalued mappings which the co-derivative has to be calculated for are typically given by normal cones N_Ω to certain closed sets Ω. For complementarity problems, for instance, $\Omega = \mathbb{R}^+_n$ and an explicit, ready for use formula for the co-derivative $D^*N_{\mathbb{R}^+_n}$ is available. In many applications, however, Ω is more complicated than just $\mathbb{R}^+_n$, for instance, it may be a general polyhedron or a set described by a finite number of smooth inequalities. Then it turns out (see [8], [7]) that, thanks to convenient calculus rules for the co-derivative, similar formulae can be obtained in those cases too, supposed that certain constraint qualifications hold true. For instance, if Ω is given by a smooth inequality system satisfying the Linear Independence Constraint Qualification (LICQ), then the co-derivative D^*N_Ω can be led back, up to an additional second order term and a linear transformation, to the well-known formula for $D^*N_{\mathbb{R}^+_n}$. In the nonregular case - if LICQ happens to be violated - still a slightly more complicated transformation formula (involving a union over non-uniquely defined multipliers) can be applied under the assumptions of Mangasarian-Fromovitz Constraint Qualification (MFCQ) and additional calmness of a certain associated multifunction (see [9]). This transformation formula holds true as an inclusion only in general, thus leaving a gap between the precise expression for the co-derivative and the one comfortably calculated from the formula. Closing this gap amounts to calculating the co-derivative 'from scratch'. Important examples, where precise formulae for the co-derivative could be obtained in the nonregular case are general polyhedra where LICQ may be violated (see [6]) and the second order cone which does not admit any description satisfying LICQ either (see [11]).

The aim of this paper is to provide explicit, ready for use co-derivative formulae for normal cone mappings to possibly nonregular inequality systems. The first part reviews some precise expressions for the co-derivative in the regular and non-regular, polyhedral setting. It is also illustrated how - similar to the reduction approach by Bonnans and Shapiro ([1], p. 240) - nonregularity of the given set Ω can be shifted in certain special situations to the simpler image set which might happen to be polyhedral and thus allow to apply the previously mentioned co-derivative formula also for certain nonregular, nonpolyhedral sets Ω. A second part of the paper deals with the transformation formula for nonregular systems mentioned above. First, it is used to derive an alternative explicit expression for the co-derivative in case of Ω being polyhedral (possibly nonregular). A couple of examples contrasts its easy use on the hand and its lack of precision on the other with an application of the previously presented precise formula. In the polyhedral case, the mentioned calmness condition required for the application of the transformation formula happens to be automatically satisfied. This is no longer true, however, and requires verification for nonlinear inequality systems. The original calmness condition is formulated for a multifunction of complicated structure involving primal and dual variables. A major part of the paper is therefore devoted to a reformulation of this condition as a constraint qualification, i.e., in terms of primal variables only. More precisely, associating with the original inequality system describing Ω the respective equality system, one has to check calmness of this equality system along with all its subsystems. A comfortable constraint qualification ensuring this property is finally derived for the special case that the number of binding inequalities exceeds the space dimension.

2 Some concepts and tools of variational analysis

We start with the definitions of the main objects in our investigation. For a closed set $\Lambda \subseteq \mathbb{R}^n$ and a point $\bar{x} \in \Lambda$, the *Fréchet normal cone* to Λ at $\bar{x} \in \Lambda$ is defined by

$$\hat{N}_\Lambda(\bar{x}) := \{x^* \in \mathbb{R}^n | \, \langle x^*, x - \bar{x} \rangle \leq o\left(\|x - \bar{x}\|\right) \quad \forall x \in \Lambda\}.$$

The *Mordukhovich normal cone* to Λ at $\bar{x} \in \Lambda$ results from the Fréchet normal cone in the following way:

$$N_\Lambda(\bar{x}) := \operatorname*{Limsup}_{x \to \bar{x}, x \in \Lambda} \hat{N}_\Lambda(x).$$

The 'Limsup' in the definition above is the upper limit of sets in the sense of Kuratowski-Painlevé, cf. [12].

For a multifunction $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$, consider a point of its graph: $(x, y) \in \operatorname{gph}\Phi$. The Mordukhovich normal cone induces the following co-derivative $D^*\Phi\,(x, y) : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ of Φ at (x, y):

$$D^*\Phi\,(x, y)\,(y^*) = \{x^* \in \mathbb{R}^n |\, (x^*, -y^*) \in N_{\operatorname{gph}\Phi}\,(x, y)\} \quad \forall y^* \in \mathbb{R}^p.$$

A multifunction $Z : Y \rightrightarrows X$ between metric spaces is said to be calm at a point $(\bar{y}, \bar{x})$ belonging to its graph, if there exist $L, \varepsilon > 0$, such that

$$d(x, Z(\bar{y})) \leq Ld(y, \bar{y}) \quad \forall x \in Z(y) \cap \mathbb{B}\,(\bar{x}, \varepsilon) \;\; \forall y \in \mathbb{B}\,(\bar{y}, \varepsilon)\,.$$

Here 'd' and '$\mathbb{B}$' refer to the distances and balls with corresponding radii in the respective metric space. For the special multifunction $Z : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^p$, defined by

$$Z\,(\alpha, \beta) := \{x \in \mathbb{R}^p | G_1(x) = \alpha,\;\; G_2(x) \leq \beta\},$$

where $G_1 : \mathbb{R}^p \to \mathbb{R}^n$ and $G_1 : \mathbb{R}^p \to \mathbb{R}^m$ are continuous mappings, it is easy to see that calmness of Z at $(0, 0, \bar{x})$ for some $\bar{x}$ satisfying $G_1\,(\bar{x}) = 0$ and $G_2\,(\bar{x}) = 0$ is equivalent with the existence of $L, \varepsilon > 0$, such that

$$d(x, Z(0, 0)) \leq L\left(\sum\nolimits_i |G_{1i}(x)| + \sum\nolimits_i \left[G_{2i}(x)\right]_+\right) \quad \forall x \in \mathbb{B}\,(\bar{x}, \varepsilon)\;. \tag{1}$$

Here, $[y]_+ := \max\{y, 0\}$.

3 On the co-derivative of normal cone mappings

3.1 Regular constraint systems

The following theorem recalls a basic transformation formula for co-derivatives which was established by Mordukhovich and Outrata in [8] (Theorem 3.4) as an inclusion with the remark that the converse inclusion is easily seen to hold as well. For the readers convenience, we've added an explicit proof in the appendix.

Theorem 3.1. *Let $C = F^{-1}(P)$, where $F : \mathbb{R}^n \to \mathbb{R}^m$ is twice continuously differentiable and $P \subseteq \mathbb{R}^m$ is some closed subset. Consider points $\bar{x} \in C$ and $\bar{v} \in N_C(\bar{x})$. If the Jacobian $\nabla F(\bar{x})$ is surjective, then*

$$D^*N_C(\bar{x}, \bar{v})(v^*) = \left(\sum_{i=1}^{m} \bar{\lambda}_i \nabla^2 F_i(\bar{x})\right) v^* + \nabla^T F\,(\bar{x})\, D^*N_P\left(F(\bar{x}), \bar{\lambda}\right)\left(\nabla F\,(\bar{x})\, v^*\right). \tag{2}$$

Here,

$$\bar{\lambda} = \left(\nabla F(\bar{x}) \nabla^T F(\bar{x})\right)^{-1} \nabla F(\bar{x})\bar{v}.$$

The value of this transformation formula relies on the fact that, starting with the co-derivative for normal cone mappings to simple objects (such as an orthant), one may pass to nonlinearly transformed constraint systems (such as differentiable inequalities). So, for instance, if

$$C = \{x \in \mathbb{R}^n \mid F_i(x) \leq 0 \quad (i = 1, \ldots, m)\},$$

where the F_i are twice continuously differentiable and $\bar{x} \in C$ satisfies the *Linear Independence Constraint Qualification*, then, putting

$$F := (F_1, \ldots, F_m)^T \quad P := \mathbb{R}^m_-,$$

one may calculate D^*N_C from $D^*N_{\mathbb{R}^m_-}$ via Theorem 3.1. To do so, one may access the following representation (see, e.g., [2], [9], [6] (Cor. 3.5)) for any $(x, v) \in \operatorname{gr} N_{\mathbb{R}^m_-}$:

$$D^*N_{\mathbb{R}^m_-}(x, v)(v^*) = \begin{cases} \emptyset & \text{if } \exists i : v_i v_i^* \neq 0 \\ \{x^* | x_i^* = 0 \ \forall i \in I_1, \ \ x_i^* \geq 0 \ \forall i \in I_2\} & \text{else} \end{cases} \tag{3}$$

where

$$I_1 := \{i | x_i < 0\} \cup \{i | v_i = 0, v_i^* < 0\}, \quad I_2 := \{i | x_i = 0, v_i = 0, v_i^* > 0\}.$$

Formula (2) is of use even in the linear case:

Corollary 3.1. *Let $C := \{x \in \mathbb{R}^n \mid Ax \leq b\}$, where $b \in \mathbb{R}^m$ and A is some matrix of order (m, n) having rank m. Then, for $\bar{x} \in C$ and $\bar{v} \in N_C(\bar{x})$, it holds that*

$$D^*N_C(\bar{x}, \bar{v})(v^*) = \begin{cases} \emptyset & \textit{if } \exists i : \bar{\lambda}_i \langle a_i, v^* \rangle \neq 0 \\ \{A^T x^* | x_i^* = 0 \ \forall i \in \tilde{I}_1, \ \ x_i^* \geq 0 \ \forall i \in \tilde{I}_2\} & \textit{else} \end{cases}$$

where

$$\begin{aligned} \tilde{I}_1 &:= \{i | \langle a_i, \bar{x} \rangle < b_i\} \cup \{i | \bar{\lambda}_i = 0, \langle a_i, v^* \rangle < 0\} \\ \tilde{I}_2 &:= \{i | \langle a_i, \bar{x} \rangle = b_i, \bar{\lambda}_i = 0, \langle a_i, v^* \rangle > 0\} \\ \bar{\lambda} &:= \left(AA^T\right)^{-1} A\bar{v} \end{aligned}$$

and the a_i refer to the rows of A.

Proof. Putting $F(x) := Ax - b$, (2) yields that

$$D^*N_C(\bar{x}, \bar{v})(v^*) = A^T D^*N_{\mathbb{R}^m_-}\left(A\bar{x} - b, \bar{\lambda}\right)(Av^*).$$

Now, the result follows from (3). □

Of course, the full-rank assumption in the corollary can be localized, so that the formula applies to any regular polyhedra defined by possibly many inequalities. Then, the matrix A has to be replaced by the submatrix corresponding to active inequalities.

3.2 Nonregular constraint systems – polyhedral image sets

Corollary 3.1 does not apply to nonregular polyhedra, for instance, it does not allow to derive a co-derivative formula for the polyhedral set $x_3 \geq \max\{|x_1|, |x_2|\}$. However, using the well-known representation of polyhedral normal cone mappings by Dontchev and Rockafellar ([2], proof of Theorem 2), one may derive an explicit co-derivative formula for arbitrary polyhedra

$$C := \{x \in \mathbb{R}^n \mid Ax \leq b\},$$

where $b \in \mathbb{R}^m$ and A is some matrix of order (m, n). To this aim, denote by a_i the rows of A and consider arbitrary $\bar{x} \in C$ and $\bar{v} \in N_C(\bar{x})$. Then, $\bar{v} = A^T \lambda$ for some $\lambda \in \mathbb{R}^m_+$. Introduce the index sets $I := \{i | \langle a_i, \bar{x} \rangle = b_i\}$ and $J := \{j | \lambda_j > 0\}$. Clearly, $J \subseteq I$. Finally, with each $I' \subseteq I$ associate its characteristic index set $\chi(I')$ consisting of those indices $j \in I$ such that for all $h \in \mathbb{R}^n$ the following implication holds true:

$$\langle a_i, h \rangle \leq 0 \quad (i \in I \setminus I'), \quad \langle a_i, h \rangle = 0 \quad (i \in I') \Longrightarrow \langle a_j, h \rangle = 0.$$

Clearly, $I' \subseteq \chi(I') \subseteq I$, $\chi(I') \subseteq \chi(I'')$ for $I' \subseteq I''$ and $I' = \chi(I')$ if the submatrix $\{a_i\}_{i \in I}$ has full rank. Now, the following relations hold true ([6], Prop. 3.2 and Cor. 3.4):

Theorem 3.2. *With the notation introduced above, one has that*

$$D^* N_C(\bar{x}, \bar{v})(v^*) = \left\{ x^* \,\middle|\, (x^*, -v^*) \in \bigcup_{J \subseteq I_1 \subseteq I_2 \subseteq I} P_{I_1,I_2} \times Q_{I_1,I_2} \right\}, \tag{4}$$

where

$$\begin{aligned} P_{I_1,I_2} &= \operatorname{con}\{a_i | i \in \chi(I_2) \setminus I_1\} + \operatorname{span}\{a_i | i \in I_1\} \\ Q_{I_1,I_2} &= \{h \in \mathbb{R}^n | \langle a_i, h \rangle = 0 \quad (i \in I_1), \quad \langle a_i, h \rangle \leq 0 \quad (i \in \chi(I_2) \setminus I_1)\} \end{aligned}$$

and 'con' and 'span' refer to the convex conic and linear hulls, respectively.

A more convenient expression avoiding the union above can be used in the following upper estimation:

$$\begin{aligned} D^* N_C(\bar{x}, \bar{v})(v^*) \subseteq\ & \operatorname{con}\{a_i | i \in \chi\left(I^a(v^*) \cup I^b(v^*)\right) \setminus I^a(v^*)\} \\ & + \operatorname{span}\{a_i | i \in I^a(v^*)\} \end{aligned} \tag{5}$$

if $\langle a_i, v^* \rangle = 0$ for all $i \in J$ and $\langle a_i, v^* \rangle \geq 0$ for all $i \in \chi(J) \setminus J$, whereas otherwise $D^* N_C(\bar{x}, \bar{v})(v^*) = \varnothing$. Here,

$$I^a(v^*) := \{i \in I | \langle a_i, v^* \rangle = 0\}, \quad I^b(v^*) := \{i \in I | \langle a_i, v^* \rangle > 0\}.$$

Corollary 3.2. $D^*N_C(\bar{x},\bar{v})(0) = \mathrm{span}\{a_i | i \in I\}$.

Proof. Since $0 \in Q_{I_1,I_2}$ for any index sets I_1, I_2, it follows from (4) that

$$D^*N_C(\bar{x},\bar{v})(0) = \bigcup_{J\subseteq I_1 \subseteq I_2 \subseteq I} P_{I_1,I_2} = P_{I,I}$$

Here, the last equality relies on the fact that $P_{I_1,I_2} \subseteq P_{I_3,I_4}$, whenever $I_1 \subseteq I_3$ and $I_2 \subseteq I_4$. □

To illustrate these characterizations, consider the following two examples:

Example 3.1. *Let* $C := Ax \leq 0$, *where*

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

Put $\bar{x} := 0$ *and* $\bar{v} := a_1 + a_2 = (0,0,2)$. *Then,* $I = \{1,2,3,4\}$, $J = \{1,2\}$ *and* $\chi(J) = I$. *Referring to (5), the condition '*$\langle a_i, v^*\rangle = 0$ *for all* $i \in J$ *and* $\langle a_i, v^*\rangle \geq 0$ *for all* $i \in \chi(J)\backslash J$*' reduces to '*$v^* = 0$*'. Moreover,* $I^a(0) = I$ *and* $I^b(0) = \varnothing$. *Consequently,* $D^*N_C(\bar{x},\bar{v})(v^*) = \varnothing$ *for* $v^* \neq 0$. *On the other hand, by Corollary 3.2,*

$$D^*N_C(\bar{x},\bar{v})(0) = \mathrm{span}\{a_i | i \in I\} = \mathrm{Im}\, A^T = \mathbb{R}^3.$$

Example 3.2. *In the previous example, put* $\bar{x} := 0$ *and* $\bar{v} := a_1 + a_3 = (-1,-1,2)$. *Then,* $I = \{1,2,3,4\}$, $J = \{1,3\}$ *and* $\chi(J) = J$. *Now, the condition '*$\langle a_i, v^*\rangle = 0$ *for all* $i \in J$ *and* $\langle a_i, v^*\rangle \geq 0$ *for all* $i \in \chi(J)\backslash J$*' reduces to* $v_1^* = v_2^* = v_3^*$. *Consequently,* $D^*N_C(\bar{x},\bar{v})(v^*) = \varnothing$ *if this last identity is violated. If it holds true, then* $D^*N_C(\bar{x},\bar{v})(0) = \mathbb{R}^3$ *by Corollary 3.2 and*

$$D^*N_C(\bar{x},\bar{v})(t,t,t) \subseteq \begin{cases} \mathrm{con}\{a_2,a_4\} + \mathrm{span}\{a_1,a_3\} & \text{if } t > 0 \\ \mathrm{span}\{a_1,a_3\} & \text{if } t < 0 \end{cases}.$$

This follows easily from (5), from the definition of A and from the already stated identity $\chi(\{1,3\}) = \{1,3\}$.

We combine the previous results for general linear and regular nonlinear constraint systems in order to calculate the co-derivative in a special nonregular, nonlinear setting. We assume that

$$C := \{x \in \mathbb{R}^n \mid AF(x) \leq b\}, \tag{6}$$

where $F : \mathbb{R}^n \to \mathbb{R}^s$ is twice continuously differentiable, $b \in \mathbb{R}^m$ and A is some matrix of order (m, s). Suppose also that $\nabla F(\bar{x})$ is surjective. Note that, in order to calculate D^*N_C, we cannot invoke Theorem 3.1 because surjectivity of $\nabla(AF)(\bar{x})$ may be violated. Nevertheless, we may rewrite the constraint set as

$$C := F^{-1}(P), \quad P := \{y \in \mathbb{R}^s \mid Ay \leq b\} \tag{7}$$

and then apply Theorem 3.1, recalling that we are able to calculate D^*N_P via Theorem 3.2. We illustrate this fact in the following example:

Example 3.3. *Let*

$$C := \{(x_1, x_2, x_3) \mid x_3 \leq -\|(x_1 + x_1^3 + x_2^4, x_1^3 + x_2 - x_2^3)\|_\infty\}$$

Evidently, C can be equivalently represented by the nonlinear inequality system

$$\begin{array}{rcl} -x_1 - x_1^3 - x_2^4 + x_3 & \leq & 0 \\ x_1 + x_1^3 + x_2^4 + x_3 & \leq & 0 \\ -x_1^3 - x_2 + x_2^3 + x_3 & \leq & 0 \\ x_1^3 + x_2 - x_2^3 + x_3 & \leq & 0 \end{array}$$

Figure 3.3 illustrates the boundary of this constraint set. At $\bar{x} = 0 \in C$, all inequal-

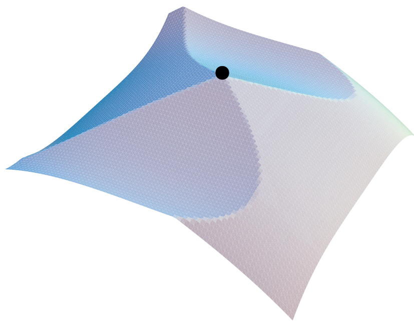

Figure 1: Illustration of the boundary of the constraint set in Example 3.3

ities are active, so their gradients cannot be linearly independent (the non-regularity can also be recognized from Fig. 3.3, where the graph exhibits four creases meeting at $\bar{x}$). This prevents an application of Theorem 3.1. However, we may write C in the form (7), where $b = 0$, A is as in Example 3.1 and

$$F(x) = (x_1 + x_1^3 + x_2^4, x_1^3 + x_2 - x_2^3, x_3)^T.$$

Evidently, $\nabla F(0) = I_3$ is surjective. As a normal vector $\bar{v} \in N_C(0)$ choose for example $\bar{v} = (-1, -1, 2)$. Because of $\nabla^2 F_i(0) = 0$ for $i = 1, 2, 3$, Theorem 3.1 provides the formula

$$D^*N_C(0, \bar{v})(v^*) = D^*N_P\left(0, \bar{v}\right)\left(v^*\right).$$

Hence, we may use for $D^*N_C(0, \bar{v})$ exactly the same estimates as obtained in Example 3.2.

3.3 Nonregular constraint systems – the use of calmness

In [9] (Th. 3.1), it was shown, how the surjectivity condition in a result like Theorem 3.1 can be weakened towards a condition, which in the setting of Theorem 3.1 would amount to the *Mangasarian-Fromovitz Constraint Qualification*, if in addition one assumes calmness of a certain multifunction. Specifying those ideas to our setting, one gets the following generalization of Theorem 3.1:

Theorem 3.3. *Consider the set* $C = \{x \in \mathbb{R}^n \mid F_i(x) \leq 0 \ (i = 1, \ldots, m)\}$*, where* $F : \mathbb{R}^n \to \mathbb{R}^m$ *is twice continuously differentiable. Fix some* $\bar{x} \in C$ *and* $\bar{v} \in N_C(\bar{x})$ *with* $F(\bar{x}) = 0$ *and suppose that the following two constraint qualifications are fulfilled:*

1. *The rows of* $\{\nabla F(\bar{x})\}$ *are positive linear independent (i.e., the Mangasarian-Fromovitz Constraint Qualification is satisfied at* $\bar{x}$*)*
2. *The multifunction*

$$M(\vartheta) := \{(x, \lambda) \mid (F(x), \lambda) + \vartheta \in \operatorname{Gr} N_{\mathbb{R}^m_-}\}$$

is calm at $(0, \bar{x}, \bar{\lambda})$ *for all* $\bar{\lambda} \in N_{\mathbb{R}^m_-}(0)$ *with* $\nabla^T F(\bar{x})\bar{\lambda} = \bar{v}$*.*

Then,

$$D^* N_C(\bar{x}, \bar{v})(v^*) \subseteq \bigcup_{\substack{\bar{\lambda} \in N_{\mathbb{R}^m_-}(F(\bar{x})) \\ \nabla^T F(\bar{x})\bar{\lambda} = \bar{v}}} \left\{ \left(\sum_{i=1}^m \bar{\lambda}_i \nabla^2 F_i(\bar{x}) \right) v^* + \nabla^T F(\bar{x}) \, D^* N_{\mathbb{R}^m_-}(0, \bar{\lambda}) \left(\nabla F(\bar{x}) \, v^* \right) \right\}.$$

As a first application of Theorem 3.3, we recover an alternative estimate of (4) and (5) in terms of dual (multipliers) rather than primal (characteristic index sets) objects.

Corollary 3.3. *Let*

$$C := \{x \in \mathbb{R}^n \mid Ax \leq b\},$$

where $b \in \mathbb{R}^m$ *and* A *is some matrix of order* (m, n)*. Fix some* $\bar{x} \in C$ *and* $\bar{v} \in N_C(\bar{x})$ *with* $A\bar{x} = b$*. If there exists some* $\xi \in \mathbb{R}^n$ *such that* $A\xi < 0$ *(component-wise), then, with the notation of Theorem 3.2,*

$$D^* N_C(\bar{x}, \bar{v})(v^*) \subseteq \operatorname{con} \{a_i | \langle a_i, v^* \rangle > 0\} + \operatorname{span} \{a_i | \langle a_i, v^* \rangle = 0\},$$

whenever there exists some $\bar{\lambda} \geq 0$ *such that* $A^T \bar{\lambda} = \bar{v}$ *and* $\bar{\lambda}_i \langle a_i, v^* \rangle = 0$ *for all* $i = 1, \ldots, m$*. Otherwise,* $D^* N_C(\bar{x}, \bar{v})(v^*) = \emptyset$*.*

Proof. In the setting of Theorem 3.3, put $F(x) := Ax - b$. Observe, that the existence of $\xi \in \mathbb{R}^n$ such that $A\xi < 0$ implies via Gordan's Theorem the first constraint qualification of the Theorem. Moreover, the multifunction considered in the second constraint qualification happens to be polyhedral, so it is calm by Robinson's well-known upper Lipschitz result for polyhedral multifunctions. Hence, one may conclude that

$$D^*N_C(\bar{x},\bar{v})(v^*) \subseteq \bigcup_{\bar{\lambda}\geq 0,\, A^T\bar{\lambda}=\bar{v}} A^T D^* N_{\mathbb{R}^m_-}\left(0,\bar{\lambda}\right)(Av^*). \tag{8}$$

From (3) we derive that the union on the right-hand side applies only if $\bar{\lambda}_i \langle a_i, v^*\rangle = 0$ for all $i = 1,\ldots,m$, in which case

$$\begin{aligned} A^T D^* N_{\mathbb{R}^m_-}\left(0,\bar{\lambda}\right)(Av^*) = \{A^T x^* | \quad & x_i^* = 0\ \forall i : \bar{\lambda}_i = 0,\ \langle a_i, v^*\rangle < 0; \\ & x_i^* \geq 0\ \forall i : \bar{\lambda}_i = 0,\ \langle a_i, v^*\rangle > 0\} \\ = \quad & \operatorname{con}\{a_i | \langle a_i, v^*\rangle > 0\} + \operatorname{span}\{a_i | \langle a_i, v^*\rangle = 0\}. \end{aligned}$$

This yields the assertion of the corollary. □

The last corollary provides a more handy formula for calculating the co-derivative of normal cone mappings to polyhedra when compared to Theorem 3.2, where characteristic index sets have to be calculated. On the other hand, it may be less precise than the latter in certain circumstances. This shall be illustrated by revisiting Examples 3.1 and 3.2:

Example 3.4 (Example 3.2 revisited)**.** *With the data from Example 3.2, the only* $\bar{\lambda}$ *with* $A^T\bar{\lambda} = \bar{v}$ *is* $\bar{\lambda} = (1,0,1,0)$. *Hence, by Corollary 3.3,*

$$D^*N_C(\bar{x},\bar{v})(v^*) \neq \varnothing \Longleftrightarrow \langle a_1, v^*\rangle = \langle a_3, v^*\rangle = 0 \Longleftrightarrow v_1^* = v_2^* = v_3^*.$$

Moreover,

$$D^*N_C(\bar{x},\bar{v})(t,t,t) \subseteq \begin{cases} \operatorname{span}\{a_1,a_2,a_3,a_4\} = \mathbb{R}^3 & \text{if } t = 0 \\ \operatorname{con}\{a_2,a_4\} + \operatorname{span}\{a_1,a_3\} & \text{if } t > 0 \\ \operatorname{span}\{a_1,a_3\} & \text{if } t < 0 \end{cases}$$

Thus, we completely recover the results of Example 3.2 obtained via Theorem 3.2.

Example 3.5 (Example 3.1 revisited)**.** *With the data from Example 3.1, there are three possibilities for* $A^T\bar{\lambda} = \bar{v}$: $\bar{\lambda} = (0,0,1,1)$, $\bar{\lambda} = (1,1,0,0)$ *and* $\bar{\lambda} = (r,r,s,s)$ *for* $r,s > 0$ *and* $r+s=1$. *Now, Corollary 3.3 implies*

$$D^*N_C(\bar{x},\bar{v})(v^*) \subseteq \begin{cases} \operatorname{span}\{a_1,a_2,a_3,a_4\} = \mathbb{R}^3 & \text{if } v^* = 0 \\ \operatorname{con}\{a_1\} + \operatorname{span}\{a_3,a_4\} & \text{if } v_2^* = v_3^* = 0,\ v_1^* < 0 \\ \operatorname{con}\{a_2\} + \operatorname{span}\{a_3,a_4\} & \text{if } v_2^* = v_3^* = 0,\ v_1^* > 0 \\ \operatorname{con}\{a_3\} + \operatorname{span}\{a_1,a_2\} & \text{if } v_1^* = v_3^* = 0,\ v_2^* < 0 \\ \operatorname{con}\{a_4\} + \operatorname{span}\{a_1,a_2\} & \text{if } v_1^* = v_3^* = 0,\ v_2^* < 0 \\ \emptyset & \text{else} \end{cases}$$

In contrast to this result, the application of Theorem 3.2 in Example 3.1 has shown that $D^*N_C(\bar{x},\bar{v})(v^*)=\emptyset$ whenever $v^*\neq 0$. In other words, the formula of Corollary 3.3 creates some additional artificial expressions in the co-derivative formula.

We now turn to an application of Theorem 3.3 in a nonlinear setting. The crucial calmness condition required there, has been investigated in [5] (see Th. 2, Th. 6, Ex. 6). As the conditions for calmness used there may be difficult to verify in general, we provide a different characterization here, where the calmness property has to be verified only for certain constraint systems in the space of x-variables which come as subsystems of the original inequality constraints. For the definition of calmness used in the following, we refer to Section 2.

Proposition 3.1. *If for all $\emptyset\neq I\subseteq\{1,\ldots,m\}$ the multifunctions*

$$H_I(\alpha)=\{x\mid F_i(x)=\alpha_i\ (i\in I),\ F_i(x)\leq 0\ (i\in I^c)\}$$

are calm at $(0,\bar{x})$, then the multifunction M introduced in Theorem 3.3 is calm at $\left(0,\bar{x},\bar{\lambda}\right)$ for any $\bar{\lambda}$ specified there.

Proof. Throughout this proof we use the $1-$norm of vectors. Note first, that for $I=\emptyset$, H_I is trivially calm as a constant multifunction. Hence, this special case can be excluded from the assumption. Next, observe that, by $F(\bar{x})=0$, one has indeed $(0,\bar{x})\in\operatorname{gr}H_I$ for all $I\subseteq\{1,\ldots,m\}$. The calmness assumption means that for any $I\subseteq\{1,\ldots,m\}$, there exist constants $\varepsilon_I,L_I>0$ such that

$$d(x,H_I(0))\leq L_I\left\|\alpha\right\|\quad\forall x\in\mathbb{B}_{\varepsilon_I}(\bar{x})\cap H_I(\alpha)\ \forall\alpha:\alpha_i\in(-\varepsilon_I,\varepsilon_I)\ \ (i\in I).$$

Putting

$$\varepsilon:=\min_{I\subseteq\{1,\ldots,m\}}\varepsilon_I,\quad L:=\max_{I\subseteq\{1,\ldots,m\}}L_I,$$

one obtains that $\varepsilon,L>0$ and

$$d(x,H_I(0))\leq L\left\|\alpha\right\|\quad\forall x\in\mathbb{B}_{\varepsilon}(\bar{x})\cap H_I(\alpha)\ \forall\alpha:\alpha_i\in(-\varepsilon,\varepsilon)\ \ (i\in I)\ \forall I\subseteq\{1,\ldots,m\}. \tag{9}$$

Due to $F(\bar{x})=0$, we may further shrink $\varepsilon>0$ such that

$$|F_i(x)|\leq\varepsilon\quad\forall x\in\mathbb{B}_{\varepsilon}(\bar{x})\ \ \forall i\in\{1,\ldots,m\}. \tag{10}$$

Now, choose arbitrary $\vartheta=(\vartheta_1,\vartheta_2)\in\mathbb{B}_{\varepsilon}(0)\times\mathbb{R}^m$ and $(x,\lambda)\in M(\vartheta)\cap(\mathbb{B}_{\varepsilon}(\bar{x})\times\mathbb{R}^m)$. We show that

$$d((x,\lambda),M(0))\leq\tilde{L}\left\|\vartheta\right\| \tag{11}$$

for some $\tilde{L}>0$. This would prove the asserted calmness of M at $\left(0,\bar{x},\bar{\lambda}\right)$ at any $\bar{\lambda}\in N_{\mathbb{R}^m_-}(0)$. Note first that $(x,\lambda)\in M(\vartheta)$ amounts to $\lambda+\vartheta_2\in N_{\mathbb{R}^m_-}(F(x)+\vartheta_1)$. Accordingly,

$$F(x)+\vartheta_1\leq 0,\ \ \lambda+\vartheta_2\geq 0,\ \ (\lambda_i+\vartheta_{2i})\left(F_i(x)+\vartheta_{1i}\right)=0\ \forall i\in\{1,\ldots,m\}. \tag{12}$$

Define

$$I' := \{i \in \{1,\ldots,m\} | F_i(x) + \vartheta_{1i} = 0 \text{ or } F_i(x) \geq 0\}.$$

Choose $\tilde{x} \in H_{I'}(0)$ such that $\|x - \tilde{x}\| = d(x, H_{I'}(0))$. Note that by definition of I', $F_i(x) < 0$ for all $i \in I'^c$. Consequently, $x \in \mathbb{B}_\varepsilon(\bar{x}) \cap H_{I'}(\alpha)$ for α defined by

$$\alpha_i := F_i(x) \quad (i \in I').$$

Since also (10) ensures that $\alpha_i \in (-\varepsilon, \varepsilon)$ for all $i \in I'$, we may apply (9) to derive that

$$d(x, H_{I'}(0)) \leq L \|\alpha\| = L \sum_{i \in I'} |F_i(x)|.$$

Now, if $i \in I'$ is such that $F_i(x) + \vartheta_{1i} = 0$, then $|F_i(x)| = |\vartheta_{1i}|$. Otherwise, by (12), $F_i(x) + \vartheta_{1i} < 0$ and, by definition of I', $F_i(x) \geq 0$. This implies $|F_i(x)| \leq |\vartheta_{1i}|$. In any case we may conclude that

$$\|x - \tilde{x}\| = d(x, H_{I'}(0)) \leq L \|\vartheta_1\|.$$

Next, define $\tilde{\lambda} \in \mathbb{R}^m$ by $\tilde{\lambda}_i := \lambda_i + \vartheta_{2i}$ if $i \in I'$ and $\tilde{\lambda}_i := 0$ if $i \in I'^c$. Then, $\tilde{\lambda} \geq 0$ by (12). Moreover, $\tilde{x} \in H_{I'}(0)$ entails that $F_i(\tilde{x}) = 0$ if $i \in I'$ and $F_i(\tilde{x}) \leq 0$ if $i \in I'^c$. In particular, $\tilde{\lambda}_i F_i(\tilde{x}) = 0$ for all $i \in \{1,\ldots,m\}$. This means that $\tilde{\lambda} \in N_{\mathbb{R}^m_-}(F(\tilde{x}))$ and, hence, $(\tilde{x}, \tilde{\lambda}) \in M(0)$. Finally observe that, for $i \in I'^c$, one has $F_i(x) + \vartheta_{1i} < 0$ and, thus, by (12), $\lambda_i = -\vartheta_{2i}$. This proves that $\tilde{\lambda} - \lambda = \vartheta_2$. Consequently,

$$\begin{aligned} d((x,\lambda), M(0)) &\leq \|(x,\lambda) - (\tilde{x},\tilde{\lambda})\| = \|x - \tilde{x}\| + \|\lambda - \tilde{\lambda}\| \\ &\leq L\|\vartheta_1\| + \|\vartheta_2\| \leq (L+1)\|\vartheta\| \end{aligned}$$

which shows (11). □

For methods to check calmness of constraint systems like those given by the multifunctions H_I in the previous proposition we refer to [4]. The next proposition shows how to get rid of inequalities for the verification of calmness in the previous proposition. More precisely, calmness has to be checked for all equality subsystems only. This proposition, which yields a slightly stronger result than needed, requires a technical lemma the proof of which is shifted to the appendix (Lemma 3.1).

Proposition 3.2. *If for all $I \subseteq \{1,\ldots,m\}$ the multifunctions*

$$\tilde{H}_I(\alpha) := \{x | F_i(x) = \alpha_i \ \ (i \in I)\}$$

are calm at $(0, \bar{x})$, then the multifunctions

$$\bar{H}_I(\alpha) = \{x | F_i(x) = \alpha_i \ \ (i \in I), \quad F_i(x) \leq \alpha_i \ \ (i \in I^c)\}$$

are also calm at $(0, \bar{x})$ for all $I \subseteq \{1,\ldots,m\}$. In particular, the multifunctions $H_I(\alpha)$ introduced in Proposition 3.1 are calm at $(0, \bar{x})$ for all $I \subseteq \{1,\ldots,m\}$.

Proof. We proceed by induction over the number m of components of F. Consider first the case $m = 1$. We either have $I = \emptyset$ or $I = \{1\}$. In the second case, one has $\bar{H}_I = \tilde{H}_I$ due to $m = 1$, hence calmness of $\bar{H}_I$ follows from that of $\tilde{H}_I$. In the first case, we apply Lemma 3.1 proved in the appendix. Referring to the notation of this lemma, we put $I^* = \emptyset$ and check the two assumptions made there. As the only set $I \subseteq \{1\}$ with $I \neq I^*$ is given by $I = \{1\}$ and then, as before, $\bar{H}_I = \tilde{H}_I$, calmness of $\bar{H}_I$ follows from that of $\tilde{H}_I$. This shows the first assumption of Lemma 3.1 to hold true. Concerning the second assumption, one has $i' = 1$ and, hence, M reduces to the trivial constant multifunction $M(\alpha, \beta) \equiv \mathbb{R}^n$ which is calm. On the other hand, the second multifunction introduced there reduces to $\bar{M} = \tilde{H}_I$, hence calmness of $\bar{M}$ follows from that of $\tilde{H}_I$. As a consequence, Lemma 3.1 yields calmness of $\bar{H}_{I^*} = \bar{H}_\emptyset$. Summarizing, the assertion of our proposition follows for the case $m = 1$. Next assume that the Proposition holds true for all $m \leq k$ and consider the case $m = k + 1$. By assumption, the $\tilde{H}_I$ are calm at $(0, \bar{x})$ for all $I \subseteq \{1, \ldots, k+1\}$. In particular, the multifunction $\bar{M}$ considered in the second assumption of Lemma 3.1 and corresponding to the case $\#I = 1$ is calm at $(0, \bar{x})$. Moreover, the induction hypothesis ensures that also the multifunctions

$$\{x | F_i(x) = \alpha_i \ \ (i \in I), \quad F_i(x) \leq \alpha_i \ \ (i \in J \diagdown I)\} \tag{13}$$

are calm at $(0, \bar{x})$ for all subsets $I \subseteq J$ and all $J \subseteq \{1, \ldots, k+1\}$ with $\#J = k$. Since the multifunction M considered in the second assumption of Lemma 3.1 is of type (13) with $J = \{1, \ldots, k+1\} \diagdown \{i'\}$, it follows that M is calm at $(0, 0, \bar{x})$. Summarizing, the second assumption of Lemma 3.1 is always satisfied no matter how the index set $I^* \subseteq \{1, \ldots, k+1\}$ is chosen in the Lemma. Therefore it is enough to check the first assumption for its application.

Now, choose an arbitrary $I^* \subseteq \{1, \ldots, k+1\}$. We have to show that $\bar{H}_{I^*}$ is calm at $(0, \bar{x})$. If $I^* = \{1, \ldots, k+1\}$, then $\bar{H}_{I^*} = \tilde{H}_{I^*}$ and calmness of $\bar{H}_{I^*}$ follows from that of $\tilde{H}_{I^*}$. If $\#I^* = k$, then the only choice for the index set I considered in the first assumption of Lemma 3.1 is $I = \{1, \ldots, k+1\}$. According to what we have shown just before, $\bar{H}_I$ is calm, so we have shown that the $\bar{H}_{I^*}$ are calm at $(0, \bar{x})$ whenever $\#I^* \geq k$. Passing to the case $\#I^* = k - 1$ and recalling that the index set I considered in the first assumption of Lemma 3.1 is always strictly larger than I^*, one derives calmness of $\bar{H}_I$ on the basis of what we have shown before due to $\#I > \#I^* = k - 1$ which amounts to $\#I \geq k$.So, the first assumption of Lemma 3.1 is satisfied again and we derive calmness of $\bar{H}_{I^*}$ whenever $\#I^* \geq k - 1$. Proceeding this way until $\#I^* = 0$, we get the desired calmness at $(0, \bar{x})$ for all subsets $I^* \subseteq \{1, \ldots, k+1\}$.

That the calmness of the $\bar{H}_I$ implies the calmness of the corresponding H_I introduced in Proposition 3.1, is an immediate consequence of the calmness definition and of the evident relations $\bar{H}_I(\alpha, 0) = H_I(\alpha)$. □

We emphasize that a result analogous to Proposition 3.2 cannot be obtained for a single constraint system (without considering subsystems). For instance, for $F(x) := (x^2, x)$ one has that the equality system $F_1(x) = \alpha_1$, $F_2(x) = \alpha_2$ is calm at $(0,0)$, whereas the inequality system $F_1(x) \leq \alpha_1$, $F_2(x) \leq \alpha_2$ is not. The reason is that subsystems need not inherit calmness (for instance, the equality subsystem $F_1(x) = \alpha_1$ fails to be calm at $(0,0)$).

We may combine Theorem 3.3, Proposition 3.1 and Proposition 3.2 to get an assumption which completely relies on constraint systems induced by F and thus can be considered to be a CQ (weaker than surjectivity) for the mapping F.

Theorem 3.4. *In the setting of Theorem 3.3 assume that*

1. *the Mangasarian Fromovitz constraint qualification is satisfied at $\bar{x}$;*
2. *all perturbed equality subsystems*

$$\{x \mid F_i(x) = \alpha_i \ (i \in I)\} \quad I \subseteq \{1, \ldots, m\}$$

are calm at $(0, \bar{x})$.

Then, the co-derivative formula of Theorem 3.3 holds true.

Remark 3.1. *If we consider the couple of constraint qualifications imposed in Theorem 3.4 as a single one and give it the name CQ*, then the following holds true for the inequality system $F(x) \leq 0$:*

$$\text{LICQ} \Longrightarrow \text{CQ}^* \Longrightarrow \text{MFCQ},$$

where MFCQ and LICQ refer to the Mangasarian-Fromovitz and Linear Independence constraint qualifications, respectively, where the latter amounts to the surjectivity condition imposed in Theorem 3.1. Indeed, the second implication being evident, suppose that $F(x) \leq 0$ satisfies LICQ at $\bar{x}$. Then, all gradients $\{\nabla F_i(\bar{x})\}_{i=1,\ldots,m}$ - and trivially all subsets of gradients - are linearly independent. But linear independence of a set of gradients implies the Aubin property and, hence, calmness for the corresponding set of equations. Consequently, CQ follows from LICQ. Summarizing, CQ* is something in between LICQ and MFCQ and it seems that it is closely related to the constant rank constraint qualification CRCQ (see [3]).*

At the end of this section, we provide a useful and easy to check constraint qualification ensuring condition 2. in Theorem 3.4.

Proposition 3.3. *In the setting of Theorem 3.3 assume that at $\bar{x}$ the following full rank constraint qualification is satisfied:*

$$\text{rank } \{\nabla F_i(\bar{x})\}_{i \in I} = \min\{n, \#I\} \quad \forall I \subseteq \{1, \ldots, m\}. \tag{14}$$

Then, the multifunctions $\tilde{H}_I$ introduced in Proposition 3.2 are calm at $(0, \bar{x})$ for all $I \subseteq \{1, \ldots, m\}$.

Proof. Choose an arbitrary $I \subseteq \{1, \ldots, m\}$. Consider first the case that $\#I \leq n$. Then, by (14), the set of gradients $\{\nabla F_i(\bar{x})\}_{i \in I}$ is linearly independent. Consequently, $\tilde{H}_I$ is calm at $(0, \bar{x})$. Now, if $\#I > n$, then select an arbitrary $J \subseteq I$ with $\#J = n$. By (14), the set of gradients $\{\nabla F_i(\bar{x})\}_{i \in I}$ is linearly independent, hence $\tilde{H}_J(0) = \{\bar{x}\}$ by the inverse function theorem. Since $F(\bar{x}) = 0$ and $\tilde{H}_I(0) \subseteq \tilde{H}_J(0)$, it follows that $\tilde{H}_I(0) = \tilde{H}_J(0)$. Moreover, according to what has been mentioned before, $\tilde{H}_J$ is calm at $(0, \bar{x})$. Consequently, there are constants $L, \varepsilon > 0$ such that

$$d(x, \tilde{H}_J(0)) \leq L \|\tilde{\alpha}\| \quad \forall x \in \tilde{H}_J(\tilde{\alpha}) \cap \mathbb{B}_\varepsilon(\bar{x}) \quad \forall \tilde{\alpha} \in \mathbb{B}_\varepsilon(0).$$

From here it follows with $\tilde{H}_I(\alpha) \subseteq \tilde{H}_J(\tilde{\alpha})$, where $\tilde{\alpha}$ is the subvector of α according to the index set $J \subseteq I$, that

$$d(x, \tilde{H}_I(0)) = d(x, \tilde{H}_J(0)) \leq L \|\tilde{\alpha}\| \leq L \|\alpha\| \quad \forall x \in \tilde{H}_I(\alpha) \cap \mathbb{B}_\varepsilon(\bar{x}) \quad \forall \alpha \in \mathbb{B}_\varepsilon(0).$$

This, however, amounts to calmness at $(0, \bar{x})$ of $\tilde{H}_I$. □

As an illustration, we revisit Example 3.3. This example being nonlinear, we cannot take assumption 2 of Theorem 3.3 automatically for granted as we could in Examples 3.4 and 3.5. On the other hand, the four constraint gradients in this example, though linearly dependent in $\mathbb{R}^3$ satisfy the full rank constraint qualification (14). Indeed, any of the 4 triples which can be selected from the original set of gradients happens to be a linearly independent set. Therefore, the second assumption of Theorem 3.4 is satisfied by virtue of Proposition 3.3. Since also the Mangasarian-Fromovitz Constraint Qualification is easily seen to be fulfilled at $\bar{x}$), we are allowed to apply the co-derivative formula of Theorem 3.3. Doing so would yield the same result as in the linearized examples discussed before.

Appendix

Proof of Theorem 3.1:

Proof. By assumption, $\nabla F(x)$ is surjective on a compact neighbourhood $\mathcal{U}(\bar{x})$. Then, by Exercise 6.7 in [12],

$$N_C(x) = \nabla^T F(x) N_P(F(x)) \quad \forall x \in \mathcal{U}(\bar{x}). \tag{15}$$

Moreover, for all $(x, \lambda, v) \in \mathcal{U}(\bar{x}) \times \mathbb{R}^s \times \mathbb{R}^n$ one has that

$$v = \nabla^T F(x)\lambda \Longrightarrow \lambda = \left(\nabla F(x) \nabla^T F(x)\right)^{-1} \nabla F(x) v. \tag{16}$$

Let $\mathcal{V}(\bar{v})$ be some compact neighbourhood of $\bar{v}$. Define π as the projection of $\mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^n$ onto its first and third components (i.e., $\pi(x, \lambda, v) = (x, v)$) and

$$D := \{(x, \lambda, v) \in \mathcal{U}(\bar{x}) \times \mathbb{R}^s \times \mathcal{V}(\bar{v}) \mid v = \nabla^T F(x)\lambda,\ \lambda \in N_P(F(x))\}.$$

It follows from (15) that

$$\operatorname{gr} N_C \cap (\mathcal{U}(\bar{x}) \times \mathbb{R}^n) = \pi(D). \tag{17}$$

Moreover, (16) implies

$$\pi^{-1}(\mathcal{U}(\bar{x}) \times \mathcal{V}(\bar{v})) \cap D \subseteq \mathcal{U}(\bar{x}) \times \left(\cup_{\substack{x \in \mathcal{U}(\bar{x}) \\ v \in \mathcal{V}(\bar{v})}} \left(\nabla F(x) \nabla^T F(x)\right)^{-1} \nabla F(x) v \right) \times \mathcal{V}(\bar{v}).$$

Obviously, the set on the right-hand side is bounded (this follows for the second factor for continuity reasons). Hence the set on the left-hand side is bounded too. Recalling that, by assumption, $(\bar{x}, \bar{v}) \in \operatorname{gr} N_C$, this allows to invoke Theorem 6.43 in [12] in order to derive from (17) the inclusion

$$\begin{aligned} N_{\operatorname{gr} N_C \cap (\mathcal{U}(\bar{x}) \times \mathbb{R}^n)}(\bar{x}, \bar{v}) &\subseteq \bigcup_{(x,\lambda,v) \in \pi^{-1}(\bar{x},\bar{v}) \cap D} \{(x^*, v^*) \mid (x^*, 0, v^*) \in N_D(x, \lambda, v)\} \\ &= \{(x^*, v^*) \mid (x^*, 0, v^*) \in N_D(\bar{x}, \bar{\lambda}, \bar{v})\}, \end{aligned}$$

where $\bar{\lambda}$ is defined in the statement of the theorem. Here, the last equality relies again on (16). On the other hand, by surjectivity of π, we may refer once more to Exercise 6.7 in [12], in order to derive from (17) that

$$N_D(\bar{x}, \bar{\lambda}, \bar{v}) = \{(x^*, 0, v^*) \mid (x^*, v^*) \in N_{\operatorname{gr} N_C \cap (\mathcal{U}(\bar{x}) \times \mathbb{R}^n)}(\bar{x}, \bar{v})\}.$$

Hence, the previous inclusion is actually an equality. We may summarize this as

$$N_{\operatorname{gr} N_C}(\bar{x}, \bar{v}) = N_{\operatorname{gr} N_C \cap (\mathcal{U}(\bar{x}) \times \mathbb{R}^n)}(\bar{x}, \bar{v}) = \{(x^*, v^*) \mid (x^*, 0, v^*) \in N_D(\bar{x}, \bar{\lambda}, \bar{v})\}. \tag{18}$$

By definition of D, we have

$$D = \Phi^{-1}(\{0\} \times \operatorname{gr} N_P) \cap (\mathcal{U}(\bar{x}) \times \mathbb{R}^s \times \mathcal{V}(\bar{v})), \tag{19}$$

where $\Phi : \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^s$ is defined as

$$\Phi(x, \lambda, v) := \left(v - \nabla^T F(x)\lambda, F(x), \lambda\right)^T.$$

Clearly, the Jacobian

$$\nabla \Phi(\bar{x}, \bar{\lambda}, \bar{v}) = \begin{pmatrix} -\sum_{i=1}^s \bar{\lambda}_i \nabla^2 F_i(\bar{x}) & -\nabla^T F(\bar{x}) & I_n \\ \nabla F(\bar{x}) & 0 & 0 \\ 0 & I_s & 0 \end{pmatrix}$$

is surjective by surjectivity of $\nabla F(\bar{x})$. This allows to employ Exercise 6.7 in [12] a third time, to see that (19) entails

$$\begin{aligned} N_D(\bar{x}, \bar{\lambda}, \bar{v}) &= N_{\Phi^{-1}(\{0\} \times \operatorname{gr} N_P)}(\bar{x}, \bar{\lambda}, \bar{v}) = \nabla^T \Phi(\bar{x}, \bar{\lambda}, \bar{v}) N_{\{0\} \times \operatorname{gr} N_P}(\Phi(\bar{x}, \bar{\lambda}, \bar{v})) \\ &= \nabla^T \Phi(\bar{x}, \bar{\lambda}, \bar{v}) \left(\mathbb{R}^n \times N_{\operatorname{gr} N_P}(F(\bar{x}), \bar{\lambda})\right). \end{aligned}$$

Given the structure of the Jacobian, we end up at

$$N_D(\bar{x},\bar{\lambda},\bar{v}) = \left\{ (x^*,\lambda^*,v^*) \;\middle|\; \begin{array}{l} x^* = -\sum_{i=1}^s \bar{\lambda}_i \nabla^2 F_i(\bar{x})v^* + \nabla^T F(\bar{x})b \\ \lambda^* = -\nabla F(\bar{x})v^* + c \\ (b,c) \in N_{\mathrm{gr}\, N_P}(F(\bar{x}),\bar{\lambda}) \end{array} \right\}.$$

Combining this with (18) and taking into account the definition of the co-derivative yields

$$N_{\mathrm{gr}\, N_C}(\bar{x},\bar{v}) = \{(x^*,v^*) \mid x^* \in -\sum_{i=1}^s \bar{\lambda}_i \nabla^2 F_i(\bar{x})v^* + \nabla^T F(\bar{x}) D^* N_P(F(\bar{x}),\bar{\lambda})(-\nabla F(\bar{x})v^*)\}.$$

Now, the assertion of the theorem follows once more from the definition of the co-derivative. □

Lemma 3.1. *Fix an arbitrary $I^* \subseteq \{1,\ldots,m\}$. Referring back to the multifunctions $\bar{H}_I$ introduced in Proposition 3.2, assume that*

1. *For all $I \neq I^*$ with $I^* \subseteq I \subseteq \{1,\ldots,m\}$ the $\bar{H}_I$ are calm at $(0,\bar{x})$.*
2. *For some $i' \in I \diagdown I^*$ the multifunctions*

$$\begin{aligned} M(\alpha,\beta) &:= \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} F_i(x) = \alpha_i \ \ (i \in I^*), \\ F_j(x) \le \beta_j \ \ (j \in \{1,\ldots,m\} \diagdown (I^* \cup \{i'\})) \end{array} \right\}, \\ \bar{M}(t) &:= \{x \in \mathbb{R}^n \,|F_{i'}(x) = t\} \end{aligned}$$

are calm at $(0,0,\bar{x})$ and $(0,\bar{x})$, respectively.

Then, $\bar{H}_{I^*}$ is calm at $(0,\bar{x})$.

Proof. Assume that $\bar{H}_{I^*}$ fails to be calm at $(0,\bar{x})$. Then, by (1), there is a sequence $x_k \to \bar{x}$ such that

$$d(x_k, \bar{H}_{I^*}(0)) > k \left(\sum_{i \in I^*} |F_i(x_k)| + \sum_{j \in \{1,\ldots,m\} \diagdown I^*} [F_j(x_k)]_+ \right). \tag{20}$$

Suppose there is some index $j' \in \{1,\ldots,m\} \diagdown I^*$ and some subsequence x_{k_l} with $F_{j'}(x_{k_l}) \ge 0$. Put $I' := I^* \cup \{j'\}$. Due to $\bar{H}_{I'}(0) \subseteq \bar{H}_{I^*}(0)$ and to $x_{k_l} \in \bar{H}_{I'}(F(x_{k_l}))$ one would arrive from (20) at

$$d(x_{k_l}, \bar{H}_{I'}(0)) > k_l \left(\sum_{i \in I'} |F_i(x_{k_l})| + \sum_{j \in \{1,\ldots,m\} \diagdown I'} [F_j(x_{k_l})]_+ \right),$$

a contradiction with assumption 1. Hence, there is some k_0 such that

$$F_j(x_k) < 0 \quad \forall k \geq k_0 \quad \forall j \in \{1, \dots, m\} \setminus I^*. \tag{21}$$

Together with (20), this implies that

$$d(x_k, \bar{H}_{I^*}(0)) > k \sum_{i \in I^*} |F_i(x_k)|. \tag{22}$$

We claim the existence of some $\rho > 0$ and $k_1 \geq k_0$ such that

$$\sum_{i \in I^*} |F_i(x_k)| > \rho |F_{i'}(x_k)| \quad \forall k \geq k_1, \tag{23}$$

where i' refers to assumption 2. Indeed, otherwise there was a subsequence x_{k_l} such that

$$\sum_{i \in I^*} |F_i(x_{k_l})| \leq l^{-1} |F_{i'}(x_{k_l})|.$$

In the following, we lead this relation to a contradiction. Now, justified by $\bar{x} \in \bar{M}(0) \neq \emptyset$, where $\bar{M}$ is defined in assumption 2, we may select for any l some $y_l \in \bar{M}(0)$ such that

$$d(x_{k_l}, \bar{M}(0)) = \|x_{k_l} - y_l\|.$$

The assumed calmness at $(0, \bar{x})$ of $\bar{M}$ entails the existence of some $L_1 > 0$ such that

$$d(x_{k_l}, \bar{M}(0)) \leq L_1 |F_{i'}(x_{k_l})| \to_l 0$$

which in turn implies that $y_l \to \bar{x}$. Consequently, for all large enough l,

$$|F_{i'}(x_{k_l})| = |F_{i'}(x_{k_l}) - F_{i'}(y_l)| \leq L_2 \|x_{k_l} - y_l\|$$

where L_2 is some Lipschitz modulus of $F_{i'}$ near $\bar{x}$. Now, referring to the multifunction M defined in assumption 2., we observe by virtue of (21) that, for all large enough l, $x_{k_l} \in M\left(\alpha^{(l)}, 0\right)$, where $\alpha_i^{(l)} := F_i(x_{k_l})$ for $i \in I^*$. Now, the assumed calmness at $(0, \bar{x})$ of M leads to

$$\begin{aligned} d(x_{k_l}, M(0,0)) &\leq L_3 \left\|\alpha^{(l)}\right\| = L_3 \sum_{i \in I^*} |F_i(x_{k_l})| \leq l^{-1} L_3 |F_{i'}(x_{k_l})| \\ &\leq l^{-1} L_3 L_2 \|x_{k_l} - y_l\| = l^{-1} L_3 L_2 d(x_{k_l}, \bar{M}(0)), \end{aligned}$$

for all large enough l. If also $l > L_3 L_2$, then

$$d(x_{k_l}, M(0,0)) < d(x_{k_l}, \bar{M}(0)). \tag{24}$$

Now, justified by $\bar{x} \in M(0,0) \neq \emptyset$, we may select $z_l \in M(0,0)$ such that

$$d(x_{k_l}, M(0,0)) = \|x_{k_l} - z_l\| \quad \forall l.$$

It follows from (24) that $z_l \notin \bar{M}(0)$, whence $F_{i'}(z_l) \neq 0$. Recalling that $F_{i'}(x_{k_l}) < 0$ for large enough l (see (21)), one would find in case of $F_{i'}(z_l) > 0$ some z' on the line segment $[x_{k_l}, z_l]$ with $F_{i'}(z') = 0$ and $\|x_{k_l} - z'\| < \|x_{k_l} - z_l\|$ yielding a contradiction with (24) due to $z' \in \bar{M}(0)$. Therefore, $F_{i'}(z_l) < 0$ and, hence, one may invoke the definition of M to infer from $z_l \in M(0,0)$ that $z_l \in \bar{H}_{I^*}(0)$ for large enough l. Now, (21) and (22) provide, for large enough l that

$$\begin{aligned} & k_l \left(\sum_{i \in I^*} |F_i(x_k)| + \sum_{j \in \{1,\ldots,m\} \setminus (I^* \cup \{i'\})} [F_j(x_{k_l})]_+ \right) \\ = \; & k_l \sum_{i \in I^*} |F_i(x_{k_l})| < d(x_{k_l}, \bar{H}_{I^*}(0)) \leq \|x_{k_l} - z_l\| = d(x_{k_l}, M(0,0)), \end{aligned}$$

a contradiction with the assumed calmness at $(0, 0, \bar{x})$ of M. This contradiction proves the desired relation (23). Using this, we may continue (22) as

$$\begin{aligned} d(x_k, \bar{H}_{I^*}(0)) &> k \left(\frac{1}{\rho+1} \sum_{i \in I^*} |F_i(x_k)| + \frac{\rho}{\rho+1} \sum_{i \in I^*} |F_i(x_k)| \right) \\ &> k \frac{\rho}{\rho+1} \left(\sum_{i \in I^* \cup \{i'\}} |F_i(x_k)| \right) \\ &= k \frac{\rho}{\rho+1} \left(\sum_{i \in I^* \cup \{i'\}} |F_i(x_k)| + \sum_{j \in \{1,\ldots,m\} \setminus (I^* \cup \{i'\})} [F_j(x_k)]_+ \right) \\ & \quad \forall k \geq k_1, \end{aligned}$$

where in the last relation, we exploited again (21). Put $I' := I^* \cup \{i'\}$. Due to $\bar{H}_{I'}(0) \subseteq \bar{H}_{I^*}(0)$ we end up at the relation

$$d(x_k, \bar{H}_{I'}(0)) > k \frac{\rho}{\rho+1} \left(\sum_{i \in I'} |F_i(x_k)| + \sum_{j \in \{1,\ldots,m\} \setminus I'} [F_j(x_k)]_+ \right) \quad \forall k \geq k_1.$$

This, however, is in contradiction with the assumed calmness at $(0, \bar{x})$ of $\bar{H}_{I'}$ (see assumption 1.). Hence, we have finally led to a contradiction our initial assumption that $\bar{H}_{I^*}$ fails to be calm at $(0, \bar{x})$. □

References

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

[2] A.L. Dontchev and R.T. Rockafellar, Characterization of strong regularity for variational inequalities over polyhedral convex sets, SIAM Journal on Optimization, 7 (1996), 1087-1105.

[3] F. Facchinei and J.-S. Pang, Finite-dimensional variational inequalities and complementarity problems, Springer, New York, 2003.

[4] J. Heerda and B. Kummer, Characterization of Calmness for Banach space mappings, Humboldt University Berlin, Institute of Mathematics, Preprint 06-26 (2006).

[5] R. Henrion and J. Outrata, Calmness of constraint systems with applications, Mathematical Programming, 104 (2005), 437-464.

[6] R. Henrion and W. Römisch, On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling, Applications of Mathematics, 52 (2007), 473-494.

[7] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory, Vol. 2: Applications, Springer, Berlin, 2006.

[8] B.S. Mordukhovich and J. Outrata, On second-order subdifferentials and their applications, SIAM Journal on Optimization, 12 (2001), 139-169.

[9] B.S. Mordukhovich and J. Outrata, Coderivative analysis of quasi-variational inequalities with applications to stability and optimization, SIAM Journal on Optimization, 18 (2007), 389-412.

[10] J.V. Outrata, A note on a class of equilibrium problems with equilibrium constraints, Kybernetika, 40 (2004), 585-594.

[11] J.V. Outrata and D. Sun, On the coderivative of the projection operator onto the second-order cone, to appear in Set-Valued Analysis.

[12] R.T. Rockafellar and R. J.-B. Wets, Variational Analysis, Springer, Berlin, 1997.

[13] J.J. Ye and X.Y. Ye, Necessary optimality conditions for optimization problems with variational inequality constraints, Mathematics of Operations Research, 22 (1997), 977-997.