Strong Lipschitz Stability of Stationary Solutions for Nonlinear Programs and Non-Polyhedral Variational Inequalies under MFCQ

Diethard Klatte[1] Bernd Kummer[2]

10 October 2003

Abstract. The stationary solution map X of a canonically perturbed nonlinear program or variational condition is studied. The focus is on characterizations for X to be locally single-valued and Lipschitz near some stationary point x^0 of an initial problem, where the Constraint Qualification MFCQ is satisfied. Since such conditions involve a non-singularity property of the strict graphical derivative TX of X, explicit formulas for TX are presented. It turns out that - even for polynomial convex problems - our stability does not only depend on certain derivatives of the problem functions at x^0. This is in contrast to various other stability concepts and holds in a similar way for the also characterized Aubin property of the same mapping. Further, we add essential examples, clarify completely the relations to Kojima's strong stability and present simplifications for linearly constrained problems, nonlinear convex programs and for the map of global minimizers as well.

Key words. Perturbed stationary solutions, nonlinear programs, variational conditions, strong Lipschitz stability, Aubin property.

AMS classification. 90C31, 49J40, 49J52, 26E25.

1 Introduction

In this paper, our basic model is the canonically perturbed nonlinear program

$$\mathrm{P}(a,b):\ \min\{f(x)-\langle a,x\rangle \mid g(x)\le b\} \tag{1}$$

with parameter $p=(a,b)\in \mathbb{R}^n\times\mathbb{R}^m$, where f and $g=(g_1,\dots,g_m)$ are defined on $\mathbb{R}^n$ and have (at least) locally Lipschitzian derivatives, briefly $f,g\in C^{1,1}$. Extensions to programs with additional parametric equations and to variational inequalities with a non-polyhedral constraint set are straightforward and will be discussed too.

Let $p\mapsto X(p)$, $p=(a,b)$, denote the map of stationary solutions x to (1) in the Karush-Kuhn-Tucker (KKT) sense. Further, let B denote the closed unit ball.

We are interested in the following local property of the multifunction $\Gamma=X$. Given a multifunction $\Gamma:\mathbb{R}^k\rightrightarrows\mathbb{R}^s$ and some $z^0\in\Gamma(0)$, Γ is said to be *strongly Lipschitz stable* at $(0,z^0)$ if

$$\begin{array}{l}\text{for certain positive } \varepsilon,\delta,L \text{ and all } p,p'\in\delta B,\\ \text{the sets } \Gamma(p)\cap(z^0+\varepsilon B) \text{ are singletons } \{z(p)\}\\ \text{and fulfill } \|z(p')-z(p)\|\le L\|p'-p\|,\end{array} \tag{2}$$

i.e., Γ is locally single-valued and Lipschitz near $(0,z^0)$. Throughout the paper, we use the following assumption concerning some initial point x^0 (without saying it explicitly)

General Assumption:
x^0 is a stationary solution of $\mathrm{P}(0)$ which satisfies the
Mangasarian-Fromovitz Constraint Qualification (MFCQ).

[1]Address: Institut für Operations Research,Universität Zürich, Moussonstrasse 15, CH-8044 Zürich, Switzerland. E–Mail: klatte@ior.unizh.ch

[2]Address: Institut für Mathematik, Humboldt–Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E–Mail: kummer@mathematik.hu-berlin.de

The purpose of the paper is to characterize strong Lipschitz stability of the stationary solution map X in terms of the given data functions and to compare the related criterion with other more or less known stability conditions.

Since we do not suppose the Linear Independence Constraint Qualification (LICQ) at the initial stationary solution x^0, we are in a situation which differs essentially from that in the standard literature.

In his landmark paper [20], Robinson particularly proved for a variational inequality

$$\text{(VI)} \quad p \in h(x) + N_C(x), \quad \text{parameter } p\,,$$

with fixed closed subset C of $\mathbb{R}^n$ that strong Lipschitz stability for solutions of some linearization of (VI) (this property was called *strong regularity* by Robinson) implies strong Lipschitz stability of solutions to (VI). This condition applies to the primal-dual solution set map X_{KKT} of a nonlinear program described by C^2 equations and inequalites, but it also implies necessarily LICQ for involved functions of type C^2 and $C^{1,1}$ [11, 13]. Strong Lipschitz stability of X_{KKT} for C^2 optimization is in fact *characterized* by both LICQ and some second order condition due to Kojima [12]. For details and bibliographical remarks in the C^2 case see, e.g., [2, 4, 7, 10, 11], for relations to the often used *coherent orientation property*, see [19, 26]. Extensions to the $C^{1,1}$ case can be found in [10, 13, 14].

Meanwhile there exists a broad literature on characterizations of stability for solutions of parametrized variational conditions and optimization problems. So it turned out that, in particular settings, strong Lipschitz stability is equivalent to other well-known properties. More precisely, it is equivalent to the Aubin property (cf. Section 4) for variational inequalities (VI) over fixed convex polyhedrons C with $h \in C^1$ [4] and both to strong stability in Kojima's [12] sense [6, 7, 11] and the Aubin property of the mapping X_{KKT} for perturbed C^2 programs [4, 10, 14].

However, a typical and crucial requirement in all these studies is LICQ or some similar non-degeneracy property of the constraints (stronger than strict MFCQ). For a recent discussion see [22, 23]. Under MFCQ alone, there is a well-known counterexample [21] of a quadratic convex program with nonlinear (non-canonical) perturbations such that the stationary solution map X is strongly stable in Kojima's sense but not strongly Lipschitz stable in the present one. All this supports the conjunction that characterizing strong Lipschitz stability needs certain new ideas.

Our approach is based on the fact that this stability of X requires a non-singularity property of its strict graphical derivative TX which, under MFCQ and the canonical perturbations in (1) and (4), will be really determined.

As a key message of this paper, it turns out that TX as well as strong Lipschitz stability of X do not only depend on the first and second derivatives of the problem functions at the base solution x^0, even for convex problems (1) only. This is in contrast to other generalized derivatives of the map X (contingent derivatives, proto-derivatives) and to other stability concepts (strong stability in Kojima's sense, upper Lipschitz properties) as well.

In Section 2 we derive a formula for TX, and our main Theorem 3.1 in Section 3 will essentially simplify this formula and hence the characterization of (non)singularity. In Section 4 we present a similar criterion for the Aubin property of the same mapping and show that both stability concepts coincide for the map X_{glob} of global solutions to (1). In both sections, we illustrate the related conditions by typical nonlinear examples. An overview over (and relations to) the mentioned other concepts will be given in Section 5 where also consequences for the cases of linear constraints as well as for non-canonical perturbations are derived.

For technical reasons which become clear below, we describe the Karush-Kuhn-Tucker points, following Kojima [12] as zeros of some function F given by the components

$$\begin{array}{rclcll} F_1(x,y) &=& Df(x) &+& \sum_{i=1}^m y_i^+ Dg_i(x), & y_i^+ = \max\{0,y_i\}, \\ F_{2,i}(x,y) &=& g_i(x) &-& y_i^-, & y_i^- = \min\{0,y_i\}. \end{array} \tag{3}$$

With the usual Lagrangian L, we have $F_1(x,y) = D_x L(x,y^+)$. For each zero of F, (x,y^+) is a usual KKT-point, and negative y_i coincide with $g_i(x)$. The KKT-points of (1) correspond to the equation $F(x,y) = p$, we put $X_{\mathsf{KKT}}(p) = \{(x,y) \mid F(x,y) = p\}$. Throughout, let

$$Y(p,x) = \{y \mid F(x,y) = p\} \quad \text{and} \quad Y^0 = Y(0,x^0).$$

As it is well-known, MFCQ at a solution $(0,x^0)$ just ensures that Y^0 is non-empty and bounded, and hence the mapping $Y(\cdot,\cdot)$ is upper semi–continuous at $(0,x^0)$.

Additional equality constraints. For simplicity of presentation, we restrict ourselves to inequality constraints, though equality constraints $g_\nu(x) = b_\nu$, $\nu = m+1,\ldots,\kappa$, can be included without affecting the results. The related Kojima function F then contains also the sum $\sum_\nu y_\nu Dg_\nu(x)$ in F_1 and the new components $F_{3,\nu} = g_\nu(x)$. In the conditions of our Theorem 3.1 then additional components α_ν and β_ν will appear where α_ν is not restricted by sign and $\beta_\nu = 0$.

Analysis of variational inequalities. As in [10], our analysis of Lipschitz stability can be applied to the solution map X of parametric variational conditions of the type

$$\mathrm{VC}(p): \quad a \in h(x) + N_{C(b)}(x), \quad \text{parameter } p = (a,b), \tag{4}$$

where $h : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz and $N_{C(b)}$ is the B-normal cone at x to the set

$$C(b) = \{x \in \mathbb{R}^n : g_i(x) \le b_i, \ i = 1,\ldots,m\}, \ g_i \in C^{1,1}(\mathbb{R}^n,\mathbb{R}).$$

Provided that $x \in C(b)$ satisfies MFCQ (or any weaker constraint qualification), it is known that x solves (4) iff $\widetilde{F}(x,y) = p$ holds for some $y \in \mathbb{R}^m$, where $\widetilde{F}$ is defined by substituting $h(x)$ for $Df(x)$ in (3) only. In our main statements, particularly when characterizing the strict graphical derivative of X and strong Lipschitz stability, *we will never make use of the special form $h = D_x f$* if nothing else is written explicitly. Therefore, this variation of F_1 allows us to deal with (1) and (4) in the same way (of course, if symmetry of Dh or the background of optimization is not needed). Again, $C^{1,1}$ equations may occur in $C(b)$ as long as MFCQ holds true.

For studying generalized derivatives in our models, it is essential that the function F in (3) has a separable product form

$$F = N(y)M(x), \quad \text{where } N(y) = (1, y^+, y^-) \in \mathbb{R}^{1+2m}$$

has (simple) directional derivatives, while $M(x)$ is a matrix of size $(1+2m, n+m)$:

$$M(x) = \begin{pmatrix} Df(x)^\mathsf{T} & g_1(x) & \ldots & g_m(x) \\ Dg_1(x)^\mathsf{T} & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ Dg_m(x)^\mathsf{T} & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & -1\,. \end{pmatrix}$$

Evidently, $M \in C^1$ holds for programs (1) if $f, g \in C^2$ and for variational conditions (4) (where h replaces Df), if $h \in C^1$, $g \in C^2$. We then speak of C^2- problems in both situations. Note that additional equality constraints simply lead us to a larger matrix and to additional (smooth) components y_ν in N.

2 Strict Graphical Derivative of the Stationary Point Map

It is well-known that strong Lipschitz stability of a (multi–)function is closely related to some kind of non-singularity of its strict graphical derivative, see, e.g., [10, 25, 13]. We recall

Definition 2.1 (*strict graphical derivative* $T\Gamma$). Let $\Gamma : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be any multifunction, $x^0 \in \Gamma(p^0)$ and $c \in \mathbb{R}^m$. Then the set $T\Gamma(p^0, x^0)(c)$ consists of all vectors $u \in \mathbb{R}^n$ such that the inclusions $x^k \in \Gamma(p^k)$ and $x^k + t_k u_k \in \Gamma(p^k + t_k c + o(t_k))$ hold for certain $t_k \downarrow 0$, $o(t_k)$ and related sequences $(x^k, u^k, p^k) \to (x^0, u, p^0)$. $\diamond$

Sets of this form have been called *strict graphical derivative* in [25], and (earlier for functions Γ) *limit sets* [27, 28], *T or Δ- derivatives* [13, 8, 10].

Definition 2.2 (*contingent derivative* $C\Gamma$). The set of all vectors $u \in \mathbb{R}^n$ such that $x^0 + t_k u_k \in \Gamma(p^0 + t_k c + o(t_k))$ holds for certain $t_k \downarrow 0$, $o(t_k)$ and $u^k \to u$, forms the *contingent derivative, also called Bouligand derivative* $C\Gamma(p^0, x^0)(c)$, cf. [1]. $\diamond$

We will be particularly interested in the case of $\Gamma(p) = X(p)$, $p = (a, b)$, $p^0 = 0$ where we may put $u^k = u$ $(\forall k)$ in the definition of TX. Indeed,

$$x^k + t_k u^k \in X(p^k + t_k c + o(t_k)) \text{ means } F(x^k + t_k u^k, y^k) = p^k + t_k c + o(t_k) \text{ for some } y^k,$$

and y^k belongs to some bounded set (by MFCQ) for (p^k, x^k) near $(0, x^0)$ and small positive t_k. Further, F is locally Lipschitz. So, replacing u^k by u leads us to

$$F(x^k + t_k u, y^k) = p^k + t_k c + o'(t_k) \tag{5}$$

with some new function o' only. Thus, with $u^k \equiv u$ one obtains the same set $TX(p^0, x^0)(c)$ (and the same $CX(p^0, x^0)(c)$, too).

Similarly, these derivatives can be applied to the function F and their factors $M(x)$ and $N(y)$. Then, as long as M is locally Lipschitz, the usual product rule of differentiation holds for these derivatives ([10]), $TM(x^0)(u)$ consists of all limits of $\mu = t^{-1}[M(x + tu) - M(x)]$, where $t = t_k \downarrow 0$ and $x = x^k \to x^0$; and TN, CN have a simple form due to the simple structure of the complementarity functions $c_i(y) = (y_i^+, y_i^-) = (y_i^+, y_i - y_i^+)$:

$$TN(y)(v) = \{(0, r, s) |\ r + s = v,\ r_i s_i \geq 0\ \forall i,\ s_i = 0 \text{ if } y_i > 0,\ r_i = 0 \text{ if } y_i < 0\}$$

$$CN(y)(v) = \{(0, r, s) \in TN(y)(v) | s_i = 0 \text{ if } y_i = 0 \text{ and } v_i \geq 0, r_i = 0 \text{ if } y_i = 0 \text{ and } v_i \leq 0\}.$$

The (full) images of these maps are polyhedral cones in $\mathbb{R}^{1+2m}$:

$$TN(y)(\mathbb{R}^m) = \{(0, \alpha, \beta) |\ \alpha_i \beta_i \geq 0\ \forall i, \beta_i = 0 \text{ if } y_i > 0,\ \alpha_i = 0 \text{ if } y_i < 0\} \tag{6}$$

$$CN(y)(\mathbb{R}^m) = \{(0, \alpha, \beta) \in TN(y)(\mathbb{R}^m) |\ \alpha_i \geq 0 \geq \beta_i \text{ if } y_i = 0\}. \tag{7}$$

Evidently, $CN(y)(v)$ consists just of the directional derivative $N'(y; v)$ of N at y in direction v; so we have $N'(y;\ \mathbb{R}^m) = CN(y)(\mathbb{R}^m)$. The set $TN(y)(\mathbb{R}^m)$ coincides with the image of Clarke's [3] generalized Jacobian $\partial N(y)$. To simplify the notation we will write

$$T_N(y) = TN(y)(\mathbb{R}^m) \quad , \quad C_N(y) = CN(y)(\mathbb{R}^m). \tag{8}$$

With (6) and (7), these derivative-images can be substituted by complementarity conditions in all subsequent formulas.

For violating strong Lipschitz stability, there are exactly two (not disjoint) possibilities:

Local unsolvability. $\exists \varepsilon > 0$ such that $X(p^k) \cap (x^0 + \varepsilon B) = \emptyset$ for certain $p^k \to 0$.

Far solutions. $\exists x^k \in X(p^k)$, $\xi^k \in X(\pi^k)$ with x^k, $\xi^k \to x^0$ and p^k, $\pi^k \to 0$ such that $x^k \neq \xi^k$ and $\|\pi^k - p^k\| / \|\xi^k - x^k\| \to 0$.

(9)

Setting, in the second case, $\xi^k = x^k + t_k u^k$, where $\|u^k\| = 1$ and $t_k > 0$, and selecting a subsequence such that $u^k \to u$, one obtains $0 \neq u \in TX(0, x^0)(a, b)$ for $(a, b) = 0$ (and vice versa). The latter is the *singular situation* we will deal with in the main part of the paper.

Definition 2.3 We speak of *singularity* if $u \in TX(0, x^0)(0)$ is valid for some $u \neq 0$. $\diamond$

For computing TX, the double limits in the definition make considerable trouble. In order to see which type of statements we may expect, let us first recall the form of CX for $C^{1,1}$ problems (1), (4), cf. [10]:

$$\begin{array}{l} \text{It holds } u \in CX(0, x^0)(a, b) \text{ iff there are } y \in Y^0 \text{ and } (\alpha, \beta) \in \mathbb{R}^{2m} \text{ satisfying} \\ a \in CF_1(., y)(x^0)u + \sum_i \alpha_i Dg_i(x^0) \ , \quad Dg(x^0)u - \beta = b \text{ and } (0, \alpha, \beta) \in C_N(y). \end{array} \tag{10}$$

The next characterization is an intermediate result and the initial point for Section 3.

Theorem 2.4 *For $C^{1,1}$ problems (1) or (4), it holds $u \in TX(0, x^0)(a, b)$ if and only if there exist some sequence $t = t_k \downarrow 0$ and assigned converging points $x \to x^0$, $y \to y^0 \in Y^0$, $y' \to y^1 \in Y^0$ (all depending on $k \to \infty$) such that*

(i) $t^{-1}[M(x + tu) - M(x)] \to M_\circ$ *for some $M_\circ \in TM(x^0)(u)$, and*

(ii) $N(y^1)M_\circ + t^{-1}[N(y') - N(y)]M(x) \to (a, b)$. $\diamond$

Before proving the theorem, we will interpret the terms and conditions appearing there.

The terms $t^{-1}[N(y') - N(y)]M(x)$ depend on Dg only and have the structure

$$\begin{array}{rcl} t^{-1} & \sum_{i=1}^m (y'^+ - y^+)_i Dg_i(x) & \text{assigned to } F_1, \\ -t^{-1} & (y'^- - y^-) & \text{assigned to } F_2. \end{array}$$

For $M \in C^1$, it follows $M_\circ = DM(x^0)u$ and

$$N(y^1)M_\circ = N(y^1)DM(x^0)u = D_x F(x^0, y^1)u = (D_x F_1(x^0, y^1)u, Dg(x^0)u) \in \mathbb{R}^{n+m}. \tag{11}$$

Generally, if M is only locally Lipschitz, the direction u is hidden in the T-derivative $M_\circ$ of M, and $M_\circ$ may depend (as a limit) on the sequence $t = t_k$ under consideration.

The point (x, y^+) is a KKT-point for problem (1),(4) with perturbation $p = F(x, y)$, while $(x', y'^+) = (x + tu, y'^+)$ is a KKT-point for (1), (4) with $p' = F(x', y')$.

The perturbation of the objective when passing from p to p' is just the linear function with gradient

$$p'_1 - p_1 = F_1(x', y') - F_1(x, y),$$

the related perturbation of the constraints becomes

$$p'_2 - p_2 = F_2(x', y') - F_2(x, y) = (g(x) - y^-) - (g(x') - y'^-).$$

The singularity condition expresses the fact that

$$\|p' - p\| << \|x' - x\| \text{ provided that } t \downarrow 0,$$

where "<<" is to read as *much smaller*. The condition (5) for TX tells us that

$$p' - p = t(a, b) + o(t).$$

These relations have to be true *for some discrete sequence $t = t_k \downarrow 0$ only.*

Proof of Theorem 2.4

($\Rightarrow$) Let $u \in TX((0,0),x^0)(a,b)$. It holds with $x' = x + tu$ and certain y', y having accumulation points $y \to y^0 \in Y^0$, $y' \to y^1 \in Y^0$ that

$$\begin{aligned} t(a,b) + o(t) &= F(x',y') - F(x,y) \\ &= N(y')M(x') - N(y)M(x) \\ &= N(y')[M(x') - M(x)] + [N(y') - N(y)]M(x). \end{aligned}$$

After division by t (and selecting an appropriate subsequence), one sees that

$$\begin{aligned} t^{-1}[M(x') - M(x)] &\to M_\circ \in TM(x^0)(u), \\ t^{-1}N(y')[M(x') - M(x)] &\to N(y^1)M_\circ. \end{aligned}$$

So the limit of

$$\Lambda = t^{-1}[N(y') - N(y)]M(x) \tag{12}$$

exists and satisfies $N(y^1)M_\circ + \Lambda \to (a,b)$, indeed.

($\Leftarrow$) Having sequences as in the theorem, it follows

$$\begin{aligned} & F(x',y') - F(x,y) \\ &= N(y')\ [M(x') - M(x)] + [N(y') - N(y)]M(x) \\ &= N(y')\ [tM_\circ + o_1(t)] + t\Lambda \\ &= N(y')\ [tM_\circ + o_1(t)] + t[(a,b) - N(y^1)\ M_\circ] + o_2(t) \\ &= t(a,b) + o_3(t), \end{aligned}$$

where o_i denote different o-type functions. This finishes the proof. $\square$

Comments.

1. By similar calculations, based on the product form of F, (simpler) formulas for the generalized derivatives $CX(0,x^0)(c)$, $TX_{\mathsf{KKT}}(0,(x^0,y^0))(c)$ and $CX_{\mathsf{KKT}}(0,(x^0,y^0))(c)$ can be derived for canonical and non-canonical parameters as well, cf. [10].

2. The hypothesis MFCQ may be weakened since, in the first part of the proof, one needs only the existence of a subsequence such that y and y' have limits in Y^0 (This is also enough for our replacement $u^k \equiv u$). In the second part, where the existence of converging y, y' is already assumed, not any constraint qualification is needed.

3. If there is a further parameter z in the functions f, g (or h) one may ask for the derivative

$$TX((0,0,0),x^0)(a,b,\zeta).$$

Now, one has $z' - z = t\,\zeta + o(t)$, and F depends on z only via $M = M(x,z)$. Writing, for $M \in C^1$, $M(x',z') = M(x,z) + DM(x,z)(x'-x, z'-z) + o(x,x',z,z')$, one sees that the term $N(y^1)\ M_\circ = N(y^1)\ DM(x^0)u$ must be replaced by the product

$$N(y^1)\ (D_xM(x^0,0)u + D_zM(x^0,0)\zeta)$$

only.

4. Difficulties for checking singularity and for deriving equivalent algebraic conditions arise from the fact that the convergence of y, y' and x may be slow in comparison to t and that the accumulation points y^0, y^1 in Y^0 may be different. $\diamond$

3 Characterizations of Singularity

3.1 Deriving the singularity condition (14)

In what follows we show that the singularity condition can be written in a way much simpler than done in Theorem 2.4. For simplicity of presentation, we assume throughout the present section that

$$M \in C^1.$$

The analysis for M being locally Lipschitz is similar, we give the corresponding modifications in several remarks.

Since we consider only points sufficiently close to $x^0 \in X(0)$, any constraint being not active at x^0 is also inactive at x near x^0 and *can be completely deleted* in our context of local stability analysis. For this reason, we also assume, without loss of generality, that

$$g(x^0) = 0.$$

Then we have $y \geq 0 \ \forall y \in Y^0$. Further, we put

$$\begin{array}{rcl} A & = & Dg(x^0) \quad \text{with rows } A_i = Dg_i(x^0)^{\mathsf{T}}, \\ Q(y) & = & D_x F_1(x^0, y). \end{array} \tag{13}$$

Theorem 3.1 *For C^2 problems (1) or (4), singularity is equivalent to the condition*

$$\begin{array}{l} \textit{There exist } u \neq 0,\ y \in Y^0 \textit{ such that for certain sequences} \\ x^k \to x^0 \quad \textit{and} \quad \alpha^k \quad \textit{with} \quad (0, \alpha^k, \beta) \in T_N(y), \quad \textit{one has} \\ Q(y)u \ + \ \sum_{i=1}^m \alpha_i^k Dg_i(x^k) \to 0 \quad \textit{and} \quad Au = \beta. \end{array} \tag{14}$$

Moreover, if $b \in ImA$, then $u \in TX(0, x^0)(a, b)$ means equivalently that certain y, x^k, α^k as in (14) satisfy $Q(y)u \ + \ \sum_{i=1}^m \alpha_i^k Dg_i(x^k) \to a$ and $Au = \beta + b$. ◇

It is important that the sequence of α^k may *diverge*, and without $b \in ImA$ we cannot pass from two elements $y^0, y^1 \in Y^0$ in Theorem 2.4 to only one, cf. Lemma 3.4.

The *persistence of local solvability* (i.e. local unsolvability as in (9) is excluded) will be studied in Section 5.3. With this respect, there are differences for the problems under consideration. In particular for (4), or for (1) with $f, g \in C^{1,1}$, such investigations need more effort and are not on the focus of this paper.

Notice, however, that without having local solvability, one obtains the local Lipschitz estimate in (2) for all existing solutions, provided that singularity does not hold.

Taking $y \geq 0 \ \forall y \in Y^0$ and formula (6) into account, we have with $Au = \beta$,

$$(0, \alpha, \beta) \in T_N(y) \quad \Leftrightarrow \quad y_i A_i u = 0 \text{ and } \alpha_i A_i u \geq 0 \ \forall i. \tag{15}$$

So (14) means, without applying any generalized derivative symbol,

$$\begin{array}{l} \text{There exist } u \neq 0,\ y \in Y^0 \text{ such that } \ y_i A_i u \ = \ 0 \ \forall i, \\ \text{and with certain sequences } x^k \to x^0 \text{ and } \alpha^k, \text{ one has} \\ \alpha_i^k A_i u \geq 0 \ \forall i \text{ and } Q(y)u + \sum_{i=1}^m \alpha_i^k Dg_i(x^k) \to 0. \end{array} \tag{16}$$

Remark 3.2 For M being locally Lipschitz, Theorem 3.1 remains true if one replaces $Q(y)u$ by any element $\phi \in TF_1(\cdot, y)](x^0)(u) \subset \mathbb{R}^n$, and if the sequence of (x^k, α^k) is assigned to vanishing $t_k \downarrow 0$ such that $t_k^{-1}[F_1(x^k + t_k u, y) - F_1(x^k, y)] \to \phi$. ◇

Theorem 3.1 will be shown by several simplifications of the condition in Theorem 2.4. The key consists in investigating the crucial limit of (12)

$$\Lambda = t^{-1}[N(y') - N(y)]M(x)$$

in condition (ii) of that theorem. In what follows we abbreviate

$$N^1 = N(y^1), \quad N^\circ = N(y^0) \ \text{ and apply that } [N^1 - N^\circ]M(x^0) = 0$$

since y^0, $y^1 \in Y^0$. Recall that $N'(y;v)$ denotes the directional derivative of N at y in direction v. Setting $y' = y^1 + v'$, $y = y^0 + v$ then, since N is piecewise linear, the formula

$$N(y') - N(y) = N^1 - N^\circ + N'(y^1;v') - N'(y^0;v)$$

holds for small norm of v' and v. With $q = N'(y^1;v') - N'(y^0;v)$ we thus obtain

$$\Lambda \ = \ t^{-1}[N^1 - N^\circ + q]M(x). \tag{17}$$

So the differences q become important for fixed y^0, y^1. These elements have the form

$$q = (0,r,s) \in \mathbb{R}^{1+2m} \ \text{ and } \ q \in D(y^1,y^0) := N'(y^1;\ \mathbb{R}^m) - N'(y^0;\ \mathbb{R}^m).$$

The set $\mathcal{D} = D(y^1,y^0)$ is a polyhedral cone. Since $y^0 \geq 0$, the images of the directional derivatives $c_i'(y_i^0;\ \mathbb{R})$ for the m functions $c_i(y_i) = (y_i^+, y_i^-)$ have the form

$$c_i'(y_i^0;\ \mathbb{R}) = \begin{cases} (\mathbb{R},0) & \text{if } y_i^0 > 0, \\ \{(\alpha,\beta) \in \mathbb{R}^2 \mid \beta \leq 0 \leq \alpha,\ \alpha\beta = 0\} & \text{if } y_i^0 = 0. \end{cases}$$

The same holds with respect to y^1. So the elements $q = (0,r,s)$ with $(r_i,s_i) \in D_i := c_i'(y_i^1;\ \mathbb{R}) - c_i'(y_i^0;\ \mathbb{R})$ build the set $\mathcal{D}$, where the sets $D_i \subset \mathbb{R}^2$ are nothing but

case 1	$D_i = \{(r_i,s_i) \mid r_i s_i \geq 0\}$	if $y_i^1 = 0,\ y_i^0 = 0$,
case 2	$D_i = \{(r_i,s_i) \mid \ s_i \leq 0\}$	if $y_i^1 = 0,\ y_i^0 > 0$,
case 3	$D_i = \{(r_i,s_i) \mid \ s_i \geq 0\}$	if $y_i^1 > 0,\ y_i^0 = 0$,
case 4	$D_i = \{(r_i,s_i) \mid \ s_i = 0\}$	if $y_i^1 > 0,\ y_i^0 > 0$.

Now we can characterize the terms Λ of (17) in more detail, where again $A_i = Dg_i(x^0)^\mathsf{T}$:

$$qM(x) \ = \ (0,\ \textstyle\sum_{i=1}^m r_i Dg_i(x)^\mathsf{T}\ ,\ -s_1,\ \ldots\ ,\ -s_m), \tag{18}$$

$$[N^1 - N^\circ]M(x) = (0,\ \textstyle\sum_{i=1}^m (y_i^1 - y_i^0) Dg_i(x)^\mathsf{T}\ ,\ ,0,\ldots,0). \tag{19}$$

Finally, recalling (11) one has

$$N^1 DM(x^0)u = D_x F(x^0,y^1)u = (\,Q(y^1)\,u\,,\ Au\,) \in \mathbb{R}^{n+m}.$$

Knowing these terms, we obtain the following corollary from Theorem 2.4.

Corollary 3.3 *It holds $u \in TX(0,x^0)(a,b)$ iff for certain y^0, $y^1 \in Y^0$ as well as some sequence $t = t_k \downarrow 0$ and assigned converging points $x \to x^0$ and $d = (d_1,\ldots,d_m) \to 0$ with $d_i = (r_i,s_i) \in D_i(y^1,y^0)$ $(\forall i)$, one has*

$$t^{-1} \sum_{i=1}^m (y_i^1 - y_i^0 + r_i) Dg_i(x) \ \to \ a - Q(y^1)u, \tag{20}$$

$$-t^{-1}s \ \to \ b - Au \quad . \tag{21}$$

◇

Simplifying the coefficients

We will finally show that the conditions of the previous corollary, for $b \in ImA$, may be reformulated by allowing to choose $y^0 = y^1$. To do this we have to deal with the sequences of t, r and s more in detail.

First we specify s_i by requiring

$$s_i = t\beta_i \text{ where } \beta = Au - b. \tag{22}$$

Let us show that this additional condition in Corollary 3.3 does not shrink the set of singular situations. Since s_i does not appear in (20), it is only important for defining the restrictions to r_i via $(r_i, s_i) \in D_i(y^1, y^0)$. Due to the cone property of the sets D_i, these restrictions remain the same after replacing s_i by any $\lambda_i s_i$ $(\lambda_i > 0)$.

If $\underline{\beta_i \neq 0}$, condition (21) yields $s_i = t\beta_i + o_i(t)$ and $\operatorname{sign}\beta_i = \operatorname{sign} s_i$. So, the setting (22) along with the requirement

$$(r_i, \beta_i) \in D_i(y^1, y^0) \tag{23}$$

leads us to the same (sign-) conditions for r_i.

If $\underline{\beta_i = 0}$ and $s_i = 0$ in the corollary then, due to the particular form of D_i, the restrictions concerning the sign of r_i disappear at all. So the same variation of r_i as before is possible.

Thus, the sequence of s may be specified by (22) for all i whereafter the conditions $(r_i, s_i) \in D_i(y^1, y^0)$ and (23) coincide.

Next we study the possible coefficients $\gamma_i := y_i^1 - y_i^0 + r_i$ in dependence of $\beta = Au - b$.

Lemma 3.4 *In the characterization of Corollary 3.3, it holds both*

(i) $\beta_i \neq 0 \;\Rightarrow\; y_i^1 = y_i^0 = 0$ *, provided that* $b \in ImA$,

(ii) $\gamma_i\ \beta_i\ \geq 0\ \ \forall i$ *if* $\|r\|$ *is small enough.* ◇

Proof. We may already suppose (22) and (23).

Let $\underline{\beta_i < 0}$. By the structure of D_i, the situation $\beta_i < 0$ is allowed in two cases. In the first case, $y_i^1 = 0$ and $y_i^0 > 0$ hold true. Then $y_i^1 - y_i^0 < 0$ and there is no sign restriction for r_i; so it follows

$$(y_i^1 - y_i^0)\beta_i > 0 \;\text{ and } \gamma_i < 0 \text{ for } |r_i| \text{ small enough.}$$

In the second case, $y_i^1 = 0 = y_i^0$ holds true. Then $r_i \leq 0$ and hence $\gamma_i \leq 0$ have to hold. Thus,

$$\beta_i < 0 \;\Rightarrow y_i^1 = 0 \;\text{ and }\; \gamma_i \leq 0.$$

Let $\underline{\beta_i > 0}$. Again, the situation $\beta_i > 0$ is allowed in two cases. In the first one, $y_i^1 > 0$ and $y_i^0 = 0$ are valid. Then $y_i^1 - y_i^0 > 0$ and there is no sign restriction for r_i. So it follows

$$(y_i^1 - y_i^0)\beta_i > 0 \;\text{ and } \gamma_i > 0 \text{ for } |r_i| \text{ small enough.}$$

In the second case, $y_i^1 = 0 = y_i^0$, it follows now $r_i \geq 0$ and $\gamma_i \geq 0$. Thus,

$$\beta_i > 0 \;\Rightarrow\; y_i^0 = 0 \;\text{ and }\; \gamma_i \geq 0.$$

Other cases are not possible. Therefore, it holds in all cases

$$\gamma_i\ \beta_i\ \geq 0 \;\text{ and }\; (y_i^1 - y_i^0)\beta_i \geq 0\ \forall\ i. \tag{24}$$

Due to $y^0, y^1 \in Y^0$, we also have $\sum_{i=1}^m (y_i^1 - y_i^0) A_i = 0$. Now suppose $b \in ImA$. Then there is some $w \in \mathbb{R}^n$ such that $\beta_i = A_i w$ and

$$\sum_{i=1}^{m} (y_i^1 - y_i^0)\beta_i = 0. \tag{25}$$

This implies that the first cases for both signs of $\beta_i \neq 0$ cannot occur. Indeed, otherwise we would obtain in view of (24), $(y_{i_0}^1 - y_{i_0}^0)\beta_{i_0} > 0$ for some i_0 and $\sum_{i=1}^m (y_i^1 - y_i^0)\beta_i > 0$, a contradiction to (25). So $\beta_i \neq 0$ is only possible for $y_i^1 = y_i^0 = 0$. This completes the proof. $\square$

Now we are able to put $y^0 = y^1$.

Corollary 3.5 *Let $b \in ImA$. Then, it holds $u \in TX(0, x^0)(a, b)$ iff there exist $y^1 \in Y^0$ satisfying $y_i^1 \beta_i = 0$ ($\forall i$) where $\beta = Au - b$, as well as some sequence $t = t_k \downarrow 0$ and assigned converging points $x \to x^0$, $(r_1, \ldots, r_m) \to 0$ such that $(r_i, \beta_i) \in D_i(y^1, y^1)$ ($\forall i$) and*

$$t^{-1} \sum_{i=1}^{m} r_i Dg_i(x) \quad \to \quad a - Q(y^1)u.$$

$\diamond$

Proof. Let $\beta = Au - b$. Then $y_i^1 \beta_i = 0$ follows from statement (i) of the Lemma. Again with $\gamma_i = (y_i^1 - y_i^0 + r_i)$, we consider any possible coefficient $\Gamma_i = t^{-1}\gamma_i$ in Corollary 3.3, and apply the equivalent conditions (22), (23) in place of (21). Setting

$$\tau = t^2 \downarrow 0 \text{ and } R_i = t\gamma_i \to 0$$

we obtain

$$\Gamma_i = t^{-1}\gamma_i = t^{-2}[t\gamma_i] = \tau^{-1} R_i.$$

If $\beta_i = 0$, it holds $(R_i, \beta_i) \in D_i(y^1, y^0)$ also for $y^0 = y^1$. If $\beta_i \neq 0$, then Lemma 3.4 yields $y_i^1 = y_i^0 = 0$. Thus $(R_i, \beta_i) \in D_i(y^1, y^1)$ is again true since $(r_i, \beta_i) \in D_i(y^1, y^1) = D_i(0, 0)$. Therefore, all coefficients Γ_i can be written in the more special form

$$\begin{array}{l} \Gamma_i = \tau^{-1} R_i, \quad \text{where } (R_i, \beta_i) \in D_i(y^1, y^1),\ R_i \to 0,\ \tau \downarrow 0 \\ \qquad\qquad\qquad \text{and} \quad y_i^1 \beta_i = 0. \end{array} \tag{26}$$

Conversely, the coefficients Γ_i defined by τ and R_i via (26) are feasible for Corollary 3.3. Identifying (r, t) and (R, τ), we have proved the assertion. $\square$

Remark 3.6 In the above proof, the more general setting $0 < \tau \leq t^2 \downarrow 0$ yields $\tau t^{-1} \leq t \downarrow 0$ and $\Gamma_i = t^{-1}(y_i^1 - y_i^0 + r_i) = \tau^{-1}[\tau t^{-1}(y_i^1 - y_i^0 + r_i)] = \tau^{-1} R_i$. Then the new scalar τ may be taken as some $t_{k'}$ having a sufficiently large index. This is important for $C^{1,1}$ problems since one may only select elements of the initial sequence t_k assigned to the element $M_\circ = (\phi, Au)^T$, ϕ from Remark 3.2. $\diamond$

Proof of Theorem 3.1

It follows directly from formula (6) that, for $y^1 \geq 0$ (and $\beta = Au - b$), the conditions

$$(r_i, \beta_i) \in D_i(y^1, y^1),\ y_i^1 \beta_i = 0\ \forall i \tag{27}$$

and

$$(0, r, \beta) \in T_N(y^1) \tag{28}$$

are equivalent. So it remains to prove that, without changing the condition, the coefficients $t^{-1}r_i$ in Corollary 3.5 can be replaced by arbitrary elements α_i such that $(0, \alpha, \beta) \in T_N(y^1)$.

Indeed, we have $(0, \alpha, \beta) \in T_N(y^1)$ for $\alpha = t^{-1}r$ since $T_N(y^1)$ is a cone and $(0, r, \beta) \in T_N(y^1)$. Conversely, every sequence of α^k satisfying $(0, \alpha^k, \beta) \in T_N(y^1)$ can be written as

$$\alpha^k = r^k/t_k \quad \text{with} \quad t_k = (k + \|\alpha^k\|^2)^{-1} \downarrow 0 \tag{29}$$

and

$$r^k = \alpha^k t_k = \alpha^k (k + \|\alpha^k\|^2)^{-1} \to 0. \tag{30}$$

Thus the sequences have still the form as required in Corollary 3.5. This proves both assertions of Theorem 3.1. $\square$

Remark 3.7 In the $C^{1,1}$ case, replace in (29) the choice of t_k by taking any $t_{k'} \le (k + \|\alpha^k\|^2)^{-1} \downarrow 0$ and replace in (30) the choice of r^k by taking any $r^{k'} = \alpha^k t_{k'} \le \alpha^k (k + \|\alpha^k\|^2)^{-1} \to 0$. $\diamond$

3.2 Discussion of the singularity condition (14)

Recall that, for C^2 problems (1) and (4), it holds for the multifunction X_{KKT} of KKT points (x, y) assigned to p and for $(x^0, y^0) \in X_{\mathsf{KKT}}(0)$:

Theorem 3.8 *X_{KKT} is strongly Lipschitz stable at $(0, (x^0, y^0))$ if and only if*

$$Au = \beta, \ Q(y^0)u + \textstyle\sum_{i=1}^m \alpha_i A_i = 0 \ \textit{and} \ (0, \alpha, \beta) \in T_N(y^0) \ \textit{imply} \ (u, \alpha, \beta) = 0. \tag{31}$$

$\diamond$

For details and generalizations to the $C^{1,1}$ case (where again $Q(y^0)u$ must be replaced by $\phi \in TF_1(\cdot, y^0)](x^0)(u)$ only) see, e.g., [10, Chapter 8]. In the next section, we shall use a similar, weaker condition, called *$TF-$ injectivity w.r. to u* in [10] (because (31) means, in fact, injectivity of $TF(x^0, y^0)$),

$$\begin{aligned} &\text{For all } y^0 \in Y^0, \\ &Au = \beta, \ Q(y^0)u + \textstyle\sum_{i=1}^m \alpha_i A_i = 0 \text{ and } (0, \alpha, \beta) \in T_N(y^0) \text{ imply } u = 0. \end{aligned} \tag{32}$$

In Section 5 we need condition (32) also with $C_N(y^0)$ at the place of $T_N(y^0)$, which we call *$CF-$ injectivity w.r. to u*. By Theorem 3.1, we observe immediately.

Corollary 3.9 *For $M \in C^1$, $TF-$ injectivity w.r. to u is necessary for strong Lipschitz stability.*

By Corollary 5.2, we will see that the reverse holds for linear constraints.

Fixing $u = 0$, condition (31) simply requires that all (active) gradients $Dg_i(x^0)$ are linearly independent, i.e., LICQ at x^0. Hence, under LICQ, strong Lipschitz stability of the stationary solution map X (being now the same as strong Lipschitz stability of X_{KKT}) depends on the first two derivatives of the data functions at x^0. The same is valid for all stability statements which (as in Section 5) can be reduced to the above injectivity conditions.

In contrast, the strong Lipschitz stability requires a "sequence characterization" (permitting even unbounded sequences) according to Theorem 3.1. The following examples show that this cannot be avoided, in general.

Example 3.10 Consider the following problem for parameter $p = 0$:

$$\min\ rx_1^2 + x_2 \ \text{ s.t. } \ g_1(x) = -x_2 \le 0,\ g_2(x) = x_1^2 - x_2 \le 0.$$

Then $Df = (2rx_1, 1)$, $Dg_1 = (0, -1)$, $Dg_2 = (2x_1, -1)$ and $x^0 = (0, 0)$ is a stationary point with $Y^0 = \{y \ge 0 \mid y_1 + y_2 = 1\}$ and $A_1 = A_2 = (0, -1)$. With $\gamma = 2r + 2y_2$, we have

$$Q(y) = \begin{pmatrix} \gamma & 0 \\ 0 & 0 \end{pmatrix}$$

Hence $u\, Q(y) = (\gamma u_1, 0)$.

Since at least one y_i is positive for $y \in Y^0$, it follows $u \perp A_i\ \forall i$ from $y_i A_i u = 0 \forall i$. Hence all u of interest have the form $u = (u_1, 0),\ u_1 \ne 0$. Condition (16) now requires exactly that for some sequence of $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ and of converging $x \to x^0$, it holds

$$(\gamma u_1, 0) + \alpha_1(0, -1) + \alpha_2(2x_1, -1) \to 0.$$

This condition cannot be satisfied *in the simple form with $x = x^0$ whenever* $\gamma \ne 0$. Note that $\gamma \ne 0$ holds for all $y \in Y^0$ if $r \notin [-1, 0]$, so convexity of the problem plays no role.

On the other hand, it suffices to put $x = (1/k,\ 0),\ \alpha_2 = -\frac{1}{2}\, k\, \gamma u_1,\ \alpha_1 = -\alpha_2$ in order to satisfy the singularity condition. So, given any r, the problem is *not strongly Lipschitz stable.*

However, if $r > 0$ then

$$\text{for each } y \in Y^0,\ Q(y) \text{ is positive definite on } K(y) = \{u \mid u \perp A_i \ \text{ if } y_i > 0\}. \qquad (33)$$

The latter is Kojima's condition [12] for *strong stability* at $(0, x^0)$, see Section 5.4. ◇

Example 3.11 Change Example 3.10 as follows

$$\min\ x_1^2 + x_2 \ \text{ s.t. } \ g_1(x) = -x_2 \le 0,\ g_2(x) = x_1^{p+1} - x_2 \le 0, \ \ p \ge 2.$$

We obtain again *singularity* at $(0, 0)$, since for any $u = (u_1, 0) \ne 0$, it holds

$$(2u_1, 0) + \alpha_1(0, -1) + \alpha_2((p + 1)x_1^p, -1) \to 0$$

for the sequences $x = (1/k, 0),\ \alpha_2 = -\frac{2u_1}{(p+1)x_1^p}$ and $\alpha_1 = -\alpha_2$.

For odd p, we are still in the class of convex problems. Nevertheless there is no possibility of identifying the singularity by using alone the first p derivatives of f and g at x^0, since these derivatives are the same for the next, strongly Lipschitz stable example with $r = 1$. ◇

Example 3.12 Change only the second constraint in Example 3.10,

$$\min\ rx_1^2 + x_2 \ \text{ s.t. } \ g_1(x) = -x_2 \le 0,\ g_2(x) = -x_2 \le 0.$$

Now the map X is *strongly Lipschitz stable* at $(0, x^0)$ for every $r \ne 0$. If $r < 0$ then the stationary points are never minimizers. This is remarkable since Kojima's strong stability holds (in the present situation) only if x^0 is a minimizer satisfying (33), cf. Section 5.4. ◇

Resume. Theorem 3.1 and the two forgoing examples show that the singularity condition for X has a new (non-) invariance quality: Even if the input functions of two problems have the same derivatives of order $k \le p < \infty$ at x^0, problem 1 may possess strongly Lipschitz stable map X, problem 2 not. This phenomenon for strong Lipschitz stability of X (and the strict graphical derivative of X) is in contrast to several other stability concepts (strong stability in Kojima's sense of X, upper Lipschitz property of X, strong regularity of the Kojima function F) and to other generalized derivatives of the mapping X (contingent derivative, proto-derivatives), see Section 5.

4 The Pseudo-Lipschitz or Aubin Property of X

Our subsequent theorem describing (the lack of) pseudo-Lipschitz behavior of $X = X(a,b)$ can be found in [10, Thm. 8.42], formulas (8.48) and (8.51). We show that this statement cannot be sharpened or simplified in a direct way, and will we apply it for establishing local solvability in Section 5.3. Further, we observe that the Aubin property coincides with strong Lipschitz stability as far as X is replaced by the map X_{glob} of all global minimizers for problem (1). We recall the definition in [1].

Definition 4.1 The map X is said to be *pseudo-Lipschitz* at $(0,x^0)$ if there are positive ε, δ and L such that $X(p) \cap (x^0 + \varepsilon B_X) \ \subset \ X(p') + L\|p'-p\|B_X \quad \forall p,p' \in \delta B_P$. $\diamond$

Other notations for the same fact are: X^{-1} is metrically regular resp. pseudo–regular [10] or X *has the Aubin property* [25]. This property has several interesting consequences and applications in the stability theory of (multi-) functions. It describes the continuity of the inverse function in the well-known Graves–Lyusternik Theorem and implies local solvability due to $x^0 \in X(0)$.

We shall say that condition (CPS) holds true if there are nontrivial pairs $(u^*, \alpha^*) \in \mathbb{R}^{n+m}$ such that, for some sequence

$$(p,x) = (p^k, x^k) \to (0, x^0)\ ,\ (p,x) \in \operatorname{gph} X, \tag{34}$$

the conditions

$$\begin{array}{lll} Dg_i(x)u^* = 0 & \text{if } i \in J^+, & \text{i.e. if } y_i > 0 \text{ for some } y \in Y(p,x), \\ \alpha_i^* \le 0 \text{ and } Dg_i(x)u^* \le 0 & \text{if } i \in J^0, & \text{i.e. if } y_i = 0 \text{ for some } y \in Y(p,x), \\ \alpha_i^* = 0 & \text{if } i \in J^-, & \text{i.e. if } y_i < 0 \text{ for some } y \in Y(p,x) \end{array} \tag{35}$$

and

$$\|u^*Q(y) + \alpha^* A\| < \varepsilon_k \downarrow 0 \quad \forall y \in Y(p,x) \tag{36}$$

are satisfied.

As usually we speak of linear constraints g_i if g_i is linear–affine.

Theorem 4.2 *([10, Thm. 8.42]) Let X belong to C^2 problems (1) or (4). Then:*
(i) The map X is not pseudo-Lipschitz at $(0,x^0)$ iff condition (CPS) is satisfied.
Further, it follows for the C^2 problem (1):
(ii) Condition (CPS) implies that (u^*, α^*, Au^*) violates (31) with some $y^0 \in Y^0$.
(iii) If all g_i are linear then (CPS) implies $u^* \neq 0$.
(iv) If all g_i are linear and (32) holds true then X is pseudo-Lipschitz at $(0,x^0)$. $\diamond$

Remark 4.3 The statements (ii), (iii) (iv) are restricted to programs (1) since they make use of the symmetry of $D^2 f$.

For $C^{1,1}$ data, the result (i) is similar: The norm (36) must be replaced by the more complicated term $\sup_{\, q \in CF_1(\cdot,y)(x)(u),\ u \in B} \langle u^*, q\rangle + \langle \alpha^*, Au\rangle$. $\diamond$

The following direct consequence of all statements (i)..(iv) presents a useful simplification:

Corollary 4.4 *For C^2 problems (1) under linear constraints, $TF-$ injectivity w.r. to u is equivalent to the Aubin property of X at $(0,x^0)$.* $\diamond$

Proof. One has only to note that $TF-$ injectivity w.r. to u (32) means that (31) cannot be violated with $y^0 \in Y^0$ and $u \neq 0$. $\square$

Comments.

1. The proof of Theorem 4.2 uses essentially the structure (10) of CX.
2. Condition (35) is equivalent to both,

$$(\alpha^*, Dg_i(x)u^*) \in T_N(y) \forall y \in Y(p,x) \quad \text{and} \quad (\alpha_i^* \leq 0 \text{ and } Dg_i(x)u^* \leq 0 \text{ if } i \in J^0).$$

Therefore, again the cone $T_N(y)$ and the matrices $Q(y) = D_x F_1(x^0, y)$, $A = Dg(x^0)$ play the crucial role, but now along with convergent sequences.

3. Due to (35) the direction u^* belongs to all tangent spaces $U(p,x)$ defined by

$$\begin{array}{ll} Dg_i(x)u \leq 0 & \text{if } g_i \text{ is active at } (p,x) \quad \text{and} \\ Dg_i(x)u = 0 & \text{if } y_i > 0 \text{ for some } y \in Y(p,x). \end{array} \tag{37}$$

Clearly, then also $Df(x)u^* = 0$ ($h(x)u^* = 0$) has to hold.

4. For sequences (p^k, x^k) (34) and related $y^k \in Y(p^k, x^k)$, convergence $y^k \to y^0 \in Y^0$ holds for some subsequence. Though the assigned index sets can switch at $(0, x^0)$ the relation $(\alpha^*, Au^*) \in T_N(y^0)$ then follows from (35) and (6); note that $T_N(.)$ is closed.
5. Obviously, the question of identifying $(p,x) \equiv (0,x^0)$ in Theorem 4.2 arises. $\diamond$

Example 4.5 To see that the latter cannot be allowed, we return to Example 3.10 where the Aubin property is violated for $r = 0$.

$$\min \ x_2 \quad \text{s.t.} \quad g_1(x) = -x_2 \leq 0, \ g_2(x) = x_1^2 - x_2 \leq 0.$$

Condition (36) requires for (p,x) from (34)

$$\|(2u_1^* y_2^+, \ 0) \ + \alpha_1^*(0, \ -1) + \alpha_2^*(0, \ -1)\| \ \leq \varepsilon_k \downarrow 0 \ \forall y \in Y(p,x). \tag{38}$$

Let $(p,x) \equiv (0,x^0)$. Since y_2 varies in $[0,1]$, (38) implies $u_1^* = 0$. Further $0 = Dg_1(x)u^* = u_2^*$ and $\alpha_i^* \leq 0$ follow from (35). But this yields $u^* = 0$ and, again due to (38), $\alpha^* = 0$. The pair (u^*, α^*) is trivial. Hence with fixed (p,x), "pseudo-singularity" cannot be seen.

May $u^* = 0$ happen in (CPS) ? Choose points of the form

$$x = (x_1, 0), \ b = (0, x_1^2), \ a = (0,0); \ x_1 \downarrow 0.$$

Then $g_1(x) = -x_2 = b_1, \ g_2(x) = x_1^2 - x_2 = b_2, \ Y(p,x) = \{(1,0)\}$ and (35) requires

$$0 = Dg_1(x)u^* = u_2^* \ \text{ and } \ \alpha_2^* \leq 0, \ Dg_2(x)u^* \leq 0.$$

Now (38) is satisfied with $u^* = 0, \alpha^* = (1,-1)$. This tells us that condition (CPS) does not necessarily imply $u^* \neq 0$.

May $\alpha^* = 0$ happen in (CPS) ? Put

$$x = (0, x_2), \ b = (-x_2, 0), \ a = (0,0); \ x_2 \downarrow 0.$$

Then $g_1(x) = -x_2 = b_1, \ g_2(x) = x_1^2 - x_2 < 0 = b_2, \ Y(p,x) = \{(1, -x_2)\}$ and (35) becomes

$$Dg_1(x)u^* = u_2^* = 0 \text{ and } \ \alpha_2^* = 0.$$

The singularity condition holds with $u^* = (1,0)$ and $\alpha^* = 0$.

We are now going to the equivalence of both regularity notions for the mapping $X_{glob}(a,b)$ *of global minimizers* to problem (1). The statement holds quite general even without supposing continuity or MFCQ.

Theorem 4.6 *For the mapping $X_{glob}(a,b)$ to programs (1), the strong Lipschitz stability and the Aubin property at $(0,x^0)$ coincide.* $\diamond$

Proof. Evidently, we have only to show that, under the Aubin property, the sets $X_{glob}(a,b)$ remain locally single valued. This is a consequence of Corollary 4.7 in [10] which states, for any b, that $X_{glob}(\cdot,b)$ is only pseudo-Lipschitz at some pair $(a,x), x \in X_{glob}(a,b)$, if $X_{glob}(a',b)$ is single–valued for all a' near a. $\square$

Thus these stability concepts for stationary points are identical in the context of convex optimization.

5 Specializations and Related Stability Concepts

In this section, we compare the characterization of strong Lipschitz stability with known conditions for several related stability notions. We also derive a specialization of the singularity condition to the case of linear constraints and give some interesting consequences. Moreover, we discuss persistence of local solvability near $(0,x^0)$ and, finally, the case of general (nonlinear) perturbations. As before, we suppose that

$$x^0 \in X(0) \text{ satisfies MFCQ and } g(x^0)=0,$$

we restrict ourselves to C^2 problems, and use the abbreviations A and $Q(y)$ as in (13).

The common point of the stability conditions for the problems (1) or (4) in the next 4 subsections consists in the fact that they depend only on the first two derivatives of f,g (resp. on Dh, Dg, D^2g) at x^0.

5.1 Singularity for Linear and Quadratic Constraints

If the constraints are sufficiently simple then strong Lipschitz stability depends again on the first two derivatives of the problem functions at x^0 only. One may even permit that one constraint is nonconvex and quadratic, the others are linear (where, under presence of equations, it does not matter which constraint is quadratic).

To show this, one has mainly to apply that, for any system of linear inequalities, the right-hand sides that permit solvability form a closed set (closed range argument). Given u define the polyhedral cone $K_u = \{\alpha \mid \alpha_i A_i u \geq 0 \ \forall i\}$.

Theorem 5.1 *Let $g_1(x) = \gamma + c^\mathsf{T}x + \frac{1}{2}x^\mathsf{T}Cx$ (with given γ, c and quadratic symmetric matrix C) and let all $g_i, i>1$ be linear. Then singularity holds if and only if there exist $u \neq 0$ and $y \in Y^0$ with $y_iA_iu = 0 \ \forall i$ such that*
(i) $Q(y)u + A^\mathsf{T}\alpha = 0$ for some $\alpha \in K_u$ or
(ii) there are $z,\zeta \in K_u, w \in \mathbb{R}^n$ with $Q(y)u + A^\mathsf{T}z + Cw = 0$, $A^\mathsf{T}\zeta = 0$ and $|\zeta_1| = 1$. $\diamond$

Proof. Given α^k and $x^k \to x^0$ according to the singularity condition of Theorem 3.1, take $Dg_1(x^k) = A_1^\mathsf{T} + C(x^k - x^0)$ into account, put $w^k = \alpha_1^k(x^k - x^0)$ and notice that

$$Dg(x^k)^\mathsf{T}\alpha^k = Cw^k + A^T\alpha^k = b^k, \quad b^k \to -Q(y)u, \ \alpha^k \in K_u. \tag{39}$$

Case 1: If $Cw^k \to 0$ (for some subsequence), the closed range argument yields that the singularity condition holds in the form $A^T\alpha = -Q(y)u$, $\alpha \in K_u$, and it follows (i).
Case 2: Otherwise, $\lambda_k := |\alpha_1^k|$ diverges, and division in (39) shows that $A^T\alpha^k/\lambda_k \to 0$.

Hence, again by the closed range argument, the equation $A^T\zeta = 0$ holds for some $\zeta \in K_u$ with $|\zeta_1| = 1$. So, (ii) is satisfied since (39) guarantees the solvability of

$$Cw + A^T z = -Q(y)u, \ z \in K_u.$$

It remains to show that, under (i) and (ii), singularity holds. Concerning (i) this is trivial, we assume (ii) and define $\alpha^k = k\zeta + z$ and $x^k = x^0 + t_k w$ with $k \to \infty$ and $t_k = (k\zeta_1 + z_1)^{-1}$. This yields

$$\alpha^k \in K_u \text{ and } Dg(x^k)^\mathsf{T}\alpha^k = A^\mathsf{T}z + (k\zeta_1 + z_1)\, t_k\, Cw = A^\mathsf{T}z + Cw = -Q(y)u.$$

Since now $x^k \to x^0$, the singularity condition is satisfied. □

It is still open whether a similar result holds for arbitrary quadratic constraint systems. If all constraints are linear then only case 1 will happen which proves:

Corollary 5.2 *Let all constraints g_i be linear for the C^2 problems (1) or (4). Then singularity holds true iff $TF-$ injectivity w.r. to u (32) is violated, i.e., there are $y \in Y^0$ and (u, α) such that $u \neq 0$, $Q(y)u + A^\mathsf{T}\alpha = 0$, $y_iA_iu = 0$ and $\alpha_iA_iu \geq 0$ $(\forall i)$.* ◇

Clearly, in this case, $Q(y)$ does not depend on y.

5.2 Upper Lipschitz Property

The following result [10, Thm. 8.24] justifies the definition of $CF-$ injectivity w.r. to u as after formula (32).

> $CF-$ injectivity w.r. to u holds true iff, for some neighborhood Ω of $x^0, L \in \mathbb{R}^+$ and sufficiently small $\|p\|$, the set $\Psi_\Omega(p) = X(p) \cap \Omega$ satisfies $\Psi_\Omega(p) \subset x^0 + L\|p\|B$.

If the latter Lipschitz estimate holds, one says that X is *locally upper Lipschitz at* $(0, x^0)$ (cf. [5, 10]).

If, additionally, $\Psi_\Omega(p)$ is not empty for sufficiently small $\|p\|$, we call X *locally upper regular at* $(0, x^0)$.

Recalling the description (7) of $C_N(y)$, $CF-$ injectivity w.r. to u can be formulated, for C^2 problems (1), in terms of a quadratic auxiliary problem and means that for all $y \in Y^0$, the origin is the unique stationary point of

$$\min\{\langle u, Q(y)u\rangle \mid u \in U^0\},$$

with $U^0 = \{u \mid Au \leq 0, u \perp Df(x^0)\}$ (standard critical cone); see, e.g., [10, §8.3].

Surprisingly, the upper Lipschitz condition becomes more complicated for the stationary point mapping $a \mapsto \hat{X}(a) := X(a, 0)$ (i.e., the constraints are not perturbed). The next statement has been originally shown via the analysis of proto-derivatives for subgradients of composed mappings (see [16, 17, 24, 25]). In the current case, the proto-derivative is just the contingent derivative. For deriving the form (40) given in the already introduced terms for C^2 problems (1), it has been supposed that $g(x^0) = 0$ and $\|Df(x^0)\|$ is sufficiently small (otherwise multiply f with some small $\lambda > 0$), see [10, Cor. 9.10].

> $\hat{X}$ is not locally upper Lipschitz at $(0, x^0)$ <u>iff</u> $CF-$ injectivity w.r. to u can be violated with some $\alpha \geq 0$ and $y^0 \in Y^0$ where, in addition, y^0 solves the LP
> $$\max_y u^\mathsf{T} D_x^2 L(x^0, y)u \ \text{ s.t. } D_xL(x^0, y) = 0 \text{ and } y_i \geq 0 \text{ if } A_iu = \alpha_i = 0, \tag{40}$$
> with optimal value 0.

Again, only the first two derivatives of f and g at x^0 are crucial.

5.3 Local Solvability

If x^0 is a local minimizer of a C^2 problem (1) at $p = 0$, then by [10, Thm. 8.36],

> $CF-$ injectivity w.r. to u implies that X is upper regular at $(0, x^0)$ and the points in $\Psi_\Omega(p)$ are again local minimizers for sufficiently small $\|p\|$.

In particular, this yields local solvability. Further, if the singularity condition of Theorem 3.1 is not satisfied then $TF-$ injectivity w.r. to u holds true. This ensures, since $C_N(y) \subset T_N(y)$, the weaker $CF-$ injectivity and in consequence:

Corollary 5.3 *For C^2 programs (1) and a local minimizer $x^0 \in X(0)$, non-singularity implies local solvability (hence also strong Lipschitz stability).* $\diamond$

For linear constraints, the situation becomes even more clear.

Theorem 5.4 *For linearly constrained C^2 programs (1), $TF-$ injectivity w.r. to u (32) coincides with the Aubin property and strong Lipschitz stability of X at $(0, x^0)$.* $\diamond$

Proof. By Corollary 5.2, the singularity condition of Theorem 3.1 is the negation of (32). Hence, under local solvability, the equivalence between strong Lipschitz stability of X and (32) is valid. On the other hand, (32) coincides with the Aubin property by Corollary 4.4. So local solvability is satisfied, too. $\square$

By the same arguments, one sees for (1) and (4) that strong Lipschitz stability just means that both singularity conditions of the Theorems 3.1 and 4.2 are not satisfied.

5.4 Strong Stability in Kojima's Sense

For a C^2 problem (1), Kojima's [12] notion of strong stability requires that, with respect to small quadratic perturbations of the objective and small right-hand-side perturbations of the constraints, the map X of stationary points is locally unique and continuous near $(0, x^0)$.

Under LICQ, the equivalence of strong Lipschitz stability and Kojima's strong stability for X is a standard fact (see, e.g., [11]). Thus let us consider the remaining case of

$$x^0 \in X(0) \ \text{ satisfies MFCQ, but not LICQ.} \tag{41}$$

For this situation, the following statement is well-known.

> Kojima [12, Thm. 7.2]: Under (41), strong stability in Kojima's sense holds true iff condition (33) is satisfied; we recall: for all $y \in Y^0$, $Q(y)$ is positive definite on the subspace $K(y) = \{u \mid u \perp A_i \ \text{ if } y_i > 0\}$.

Obviously, this implies that x^0 is a local minimizer for (1) at $p = 0$ and, in addition, that one cannot find $y \in Y^0$ and (u, α) with $u \neq 0$ such that

$$Q(y)u + \sum_{i=1}^{m} \alpha_i A_i = 0, \quad y_i A_i u = 0 \text{ and } \alpha_i A_i u \geq 0 \ (\forall i)$$

(otherwise multiply the first equation with u). The latter is $TF-$ injectivity w.r. to u. So, we arrived at the following interrelation.

Corollary 5.5 *Let $x^0 \in X(0)$ satisfy (41).*
(i) Under linear constraints, Kojima's strong stability implies strong Lipschitz stability.
(ii) If x^0 is a local minimizer for $p = 0$, then strong Lipschitz stability implies Kojima's strong stability. $\diamond$

Proof. Statement (i) now follows from Theorem 5.4, statement (ii) from [10, Cor. 8.37], which just asserts that Kojima's condition is satisfied. $\square$

Due to Example 3.12, statement (i) cannot be reversed, due to Example 3.10, the implication (i) is not valid for nonlinear convex problems.

5.5 Consequences for Nonlinear Perturbations

Consider the parametric program:

$$\text{(P)}(z,p): \quad \begin{array}{l} \min_x \ f(x,z) - a^\mathsf{T} x \\ \text{s.t.} \ \ g_i(x,z) \le b_i \ (i = 1, \dots, m), \ p = (a,b), \end{array}$$

and suppose that $f, g \in C^2$ with respect to (x,z) and $D^2_{xz} f$ and $D^2_{xz} g$ are locally Lipschitz. Let $(z,p) \mapsto \widetilde{X}(z,p)$ denote the stationary point map. Then it is well-known that strong stability in Kojima's sense and the upper Lipschitz property of X carry over to the mapping $\widetilde{X}$.

On the other hand, Robinson [21, p. 219] gave an example for a simple problem, namely, to minimize $\sum_{i=1}^m x_i^2$ s.t. nonlinearly (by z) perturbed linear constraints satisfying MFCQ, with two important properties: Under canonical perturbations of the problem to $z = 0$, X is strongly Lipschitz stable at $(0, x^0)$. In contrast, $\widetilde{X}$ is not strongly Lipschitz stable. So there is a gap between canonical and nonlinear perturbations.

We explain this gap in the context of a parametric variational inequality, where $h = h(x,z)$ and $g = g(x,z)$ are sufficiently smooth. Our construction is standard in linking an inverse function theorem with an implicit function theorem. Under MFCQ and suitable second-order conditions, it has been also successfully used with respect to upper Lipschitz behavior, Kojima's strong stability and calculation of the contingent derivative of $\widetilde{X}$, see e.g. [9, 10, 12, 21]. However, strong Lipschitz stability requires additional regularity assumptions like a constant rank constraint qualification [18].

Due to existence and continuity of the related derivatives, we have

$$\begin{aligned} h(x,z) &= h(x,0) + D_z h(x,0) z + o_1(x,z) \\ g(x,z) &= g(x,0) + D_z g(x,0) z + o_2(x,z) \\ D_x g(x,z) &= D_x g(x,0) + D_z D_x g(x,0) z + o_3(x,z), \end{aligned}$$

and we obtain for any $x \in X(z)$ from

$$h(x,z) + y^{+T} D_x g(x,z) = 0 \ \text{ and } \ g(x,z) = y^-$$

that

$$h(x,0) + D_z h(x,0) z + o_1 \ + \ y^{+T}(D_x g(x,0) + D_z D_x g(x,0) z + o_3) = 0$$

and

$$g(x,0) + D_z g(x,0) z + o_2 = y^-.$$

Here, $y \in Y(z,x)$ is some (Kojima) dual vector to x for parameter z. Setting

$$a = D_z h(x,0) z + o_1 + \ y^{+T}(D_z D_x g(x,0) z + o_3) \ \text{ and } \ b = -(D_z g(x,0) z + o_2), \tag{42}$$

we also know that x is stationary for the canonically perturbed problem

$$a + h(x, 0) \in N_{C(b)} \quad \text{where } C(b) \text{ is given by } g(x, 0) \le b.$$

For small $\|z\|$ and x near x^0, the duals y are locally bounded (MFCQ) and close to Y^0. So the linear maps in (42)

$$D_z h(x, 0) \ , \ y^{+T} D_z D_x g(x, 0) \quad \text{and} \quad D_z g(x, 0)$$

are uniformly Lipschitz (rank L') in z. The same holds true (uniformly for small $\|z\|$ and x near x^0) for the functions o_k under consideration. This ensures, for $x \in X(z)$ near x^0, a (pointwise) Lipschitz estimate

$$\|x - x^0\| \le \ L(\|a\| + \|b\|) \le \ LL'\|z\|.$$

However, for a second pair of points $x' \in X(z')$ and $y' \in Y(z', x')$, the crucial terms

$$y^{+T} D_z D_x g(x, 0) \quad \text{and} \quad y'^{+T}(D_z D_x g(x', 0)$$

of (42) are (in general) only close to each other if $\|y' - y\|$ is sufficiently small. Only in this case, one may estimate (with new L')

$$\|x' - x\| \le \ L(\|a' - a\| + \|b' - b\|) \le \ LL'\|z' - z\|.$$

Hence, for extending strong Lipschitz stability to arbitrary nonlinear parametrizations it is enough (under the necessary smoothness) to require the so-called *strict Mangasarian-Fromovitz Constraint Qualification* which just ensures that Y^0 is single-valued.

References

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

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

[3] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[4] A. Dontchev and R.T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM Journal on Optimization*, 6:1087–1105, 1996.

[5] A. Dontchev and R.T. Rockafellar. Characterizations of Lipschitz stability in nonlinear programming. In A.V. Fiacco, editor, *Mathematical Programming with Data Perturbations*, pages 65–82, Marcel Dekker, 1998.

[6] H.Th. Jongen, D. Klatte and K. Tammer. Implicit functions and sensitivity of stationary points. *Mathematical Programming*, 49:123–138, 1990.

[7] D. Klatte and B. Kummer. Strong stability in nonlinear programming revisited. *Journal of the Australian Mathematical Society, Series B*, 40:336–352, 1999.

[8] D. Klatte and B. Kummer. Generalized Kojima functions and Lipschitz stability of critical points. *Computational Optimization and Applications*, 13:61–85, 1999.

[9] D. Klatte and B. Kummer. Contingent derivatives of implicit (multi–) functions and stationary points. *Annals of Operations Research*, 101:313–331, 2001.

[10] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization - Regularity, Calculus, Methods and Applications.* Nonconvex Optimization and Its Applications. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.

[11] D. Klatte and K. Tammer. Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization. *Annals of Operations Research*, 27:285–307, 1990.

[12] M. Kojima. Strongly stable stationary solutions in nonlinear programs. In S.M. Robinson, editor, *Analysis and Computation of Fixed Points*, pages 93–138. Academic Press, New York, 1980.

[13] B. Kummer. Lipschitzian inverse functions, directional derivatives and application in $C^{1,1}$ optimization. *Journal of Optimization Theory and Applications*, 70:559–580, 1991.

[14] B. Kummer. Lipschitzian and pseudo-Lipschitzian inverse functions and applications to nonlinear programming. In A.V. Fiacco, editor, *Mathematical Programming with Data Perturbations*, pages 201–222. Marcel Dekker, New York, 1998.

[15] A.B. Levy, R.A. Poliquin and R.T. Rockafellar. Stability of locally optimal solutions. *SIAM Journal on Optimization*, 10:580–604, 2000.

[16] A.B. Levy and R.T. Rockafellar. Sensitivity of solutions in nonlinear programs with nonunique multipliers. In D.-Z. Du, L. Qi, and R.S. Womersley, editors, *Recent Advances in Nonsmooth Optimization*, pages 215–223. World Scientific Press, Singapore, 1995.

[17] R.A. Poliquin. Proto-differentiation of subgradient set-valued mappings. *Canadian Journal of Mathematics*, XLII, No.3:520–532, 1990.

[18] D. Ralph and S. Dempe. Directional derivatives of the solution of a parametric nonlinear program. *Mathematical Programming*, 70:159–172, 1995.

[19] D. Ralph and S. Scholtes. Sensivitity analysis of composite piecewise smooth equations. *Mathematical Programming, Series B*, 76:593–612, 1997.

[20] S.M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5:43–62, 1980.

[21] S.M. Robinson. Generalized equations and their solutions. Part II: Applications to nonlinear programming. *Mathematical Programming Study*, 19:200–221, 1982.

[22] S.M. Robinson. Constraint nondegeneracy in variational analysis. *Mathematics of Operations Research*, 28:201–232, 2003.

[23] S.M. Robinson. Variational conditions with smooth constraints: structure and analysis. *Mathematical Programming*, 97: 245–265, 2003.

[24] R.T. Rockafellar. First and second order epi-differentiability in nonlinear programming. *Transactions of the American Mathematical Society*, 207:75–108, 1988.

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

[26] S. Scholtes. Introduction to Piecewise Differentiable Equations. *Institut für Statistik und Mathematische Wirtschaftstheorie, Preprint No. 53*, Universität Karlsruhe, 1994.

[27] L. Thibault. Subdifferentials of compactly Lipschitz vector-valued functions. *Annali di Matematica Pura ed Applicata*, 4:157–192, 1980.

[28] L. Thibault. On generalized differentials and subdifferentials of Lipschitz vector-valued functions. *Nonlinear Analysis: Theory, Methods & Applications*, 6:1037–1053, 1982.