Preprint manuscript No.
(will be inserted by the editor)

Aubin Property and Uniqueness of Solutions in Cone Constrained Optimization

Diethard Klatte · **Bernd Kummer**

submitted 6 January 2012, revised 11 June 2012

Abstract We discuss conditions for the Aubin property of solutions to perturbed cone constrained programs, by using and refining results given in [10]. In particular, we show that constraint nondegeneracy and hence uniqueness of the multiplier is necessary for the Aubin property of the critical point map. Moreover, we give conditions under which the critical point map has the Aubin property if and only if it is locally single-valued and Lipschitz.

Keywords Cone constrained optimization · Aubin property · Critical points · Constraint nondegeneracy · Locally single-valued solutions

Mathematics Subject Classification (2000) 49K40 · 90C31 · 49J53 · 90C22

1 Introduction

In this paper we consider the canonically perturbed cone constrained program

$$P(p), p = (a,b): \qquad \min_x \; f(x) - \langle a,x\rangle \quad \text{subject to} \quad g(x) - b \in K, \qquad (1)$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $K \subset \mathbb{R}^m$ is a closed convex set, $\langle \cdot, \cdot \rangle$ is the Euclidean inner product (with induced norm $\|\cdot\|$), and the parameter vector $p = (a,b)$ varies near the origin. Put (P) = $P(0)$. Note that the results of this paper remain true if f and g also vary in a suitable way in some function space or by certain parameterizations but we avoid this (i) because of size restrictions for this paper and (ii) for an intrinsic reason: the stability characterizations given below depend crucially on the canonical perturbations $f - \langle a, \cdot \rangle$ and $g - b$.

Diethard Klatte
Institut für Betriebswirtschaftslehre, Professur Mathematik für Ökonomen, Universität Zürich, Moussonstr. 15, CH-8044 Zürich, Switzerland. E-mail: diethard.klatte@business.uzh.ch

Bernd Kummer
Institut für Mathematik, Humboldt–Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany. E-mail: kummer@math.hu-berlin.de

We speak here of *cone constraints*, since K is often a cone in applications. Note that K may be a set of symmetric matrices. In this case, the standard reformulation of a symmetric (d,d)-matrix A as a vector $\mathrm{svec}(A) \in \mathbb{R}^{d(d+1)/2}$ leads to a problem of type (1). Thus, problem (1) particularly covers standard nonlinear programs, second-order cone programs and semi-definite programs under perturbations.

The *Lagrange function* of problem (1) is given by

$$L(x,\lambda) = f(x) + \langle \lambda, g(x) \rangle.$$

If f and g are C^1 functions, and $\bar{x}$ is a feasible point of $P(p)$ satisfying a constraint qualification, then there is some λ such that $(\bar{x},\lambda)$ fulfils the *Karush-Kuhn-Tucker (KKT) conditions*

$$D_x L(\bar{x},\lambda) = a\,, \quad \lambda \in N_K(g(\bar{x}) - b), \tag{2}$$

where $N_K(\bar{y}) = \{z \in \mathbb{R}^m \,|\, \langle z, y - \bar{y} \rangle \le 0 \ \forall y \in K\}$ is the *normal cone* of K at $\bar{y} \in K$. Note that by definition $N_K(\bar{y}) = \emptyset$ if $\bar{y} \notin K$. In this sense, the feasibility condition $g(\bar{x}) - b \in K$ is included when saying that $(\bar{x},\lambda)$ fulfils (2). For first-order optimality conditions and the theory of cone constrained programs at all we refer e.g. to Bonnans and Shapiro [1].

For $p = (a,b)$, denote by $\Sigma(p)$ the set of all solutions $(\bar{x},\lambda)$ to the system (2). A point $(\bar{x},\lambda) \in \Sigma(p)$ will be called *critical point* of problem $P(p)$. By $\Lambda(\bar{x},p) = \{\lambda \,|\, (\bar{x},\lambda) \in \Sigma(p)\}$ we denote the set of *multipliers* associated with some $\bar{x}$ in the set $S(p) = \{x \,|\, \exists \lambda :\ (x,\lambda) \in \Sigma(p)\}$ of *stationary solutions* of $P(p)$.

The focus of our paper is to study situations in which the critical point *multifunction* Σ (resp. Λ or S) is in fact a locally single-valued and Lipschitz *function*, provided it has some multi-valued Lipschitz behavior called Aubin property. We say that a multifunction Γ from $\mathbb{R}^d$ to $\mathbb{R}^s$ has the *Aubin property* at some point $(\bar{p},\bar{z}) \in \mathrm{gph}\,\Gamma = \{(p,z) \,|\, z \in \Gamma(p)\}$ (or, synonymously, the inverse multifunction Γ^{-1} is *metrically regular* at $(\bar{z},\bar{p})$), if there are neighborhoods U of $\bar{p}$ and V of $\bar{z}$ as well as some constant $c > 0$ such that for all $p, p' \in U$ and all $z \in \Gamma(p) \cap V$

$$\text{there exists some } z' \in \Gamma(p') \text{ such that } \|z - z'\| \le c\|p - p'\|.$$

It is immediate from the definition that the Aubin property of Γ at $(\bar{p},\bar{z})$ implies that Γ has the Aubin property also at $(p,z) \in \mathrm{gph}\,\Gamma$ near $(\bar{p},\bar{z})$. If $\Gamma(p) \cap V$ is single-valued on U then it becomes a locally Lipschitz function under the Aubin property. In this case, Γ is called *locally single-valued and Lipschitz* around $(\bar{p},\bar{z})$, or, synonymously, Γ^{-1} is *strong regular* at $(\bar{z},\bar{p})$. For a detailed discussion of relations between different stability and regularity notions we refer, e.g., to [17,10,6].

If $F : \mathbb{R}^d \to \mathbb{R}^d$ is a locally Lipschitz and directionally differentiable function, then the Aubin property of $\Gamma = F^{-1}$ at $(\bar{z},\bar{p})$ gives that $\bar{z}$ is an isolated (i.e., locally unique) solution of the equation $F(z) = \bar{p}$, see Fusek [7]. Of course, this applies to the critical point map $\Sigma(a) = (Df)^{-1}(a)$ of the unconstrained problem (1) for $F = Df$. Unfortunately, this does not mean that Df^{-1} is locally single-valued in this case, see Kummer's example of a piecewise quadratic C^1 function f in [12, Ex. 3] (cf. also [10, Ex. BE.4]). In contrast, if f is a C^2 function then Df^{-1} becomes locally single-valued under the Aubin property, by standard calculus arguments.

What about constrained problems? For *global minimizers*, uniqueness follows from the Aubin property for general optimization problems, we refer to [10] and section 4 below. For *critical points*, the situation is much more involved. Dontchev and Rockafellar [5] studied a variational inequality $F(z,t) - p \in N_C(z)$ for a polyhedral convex set C and a C^1 function $F : \mathbb{R}^d \times \mathbb{R}^\kappa \to \mathbb{R}^d$ with perturbation (t,p). They proved the equivalence of metric and strong regularity at solutions for this model, see Proposition 1 below. This result applies to the critical point map of a perturbed nonlinear program with C^2 data, i.e., in our context, to the problem (1) with convex polyhedral set K and $f,g \in C^2$.

Kummer [12] (see also [10, §7]) extended this results to the solution map of a so-called generalized Kojima system (cf. section 5 below) which covers the Dontchev-Rockafellar model. In a recent paper, Outrata and Ramírez [14] show the equivalence of the Aubin property and strong regularity for the stationary solution set map of a second-order cone program under constraint nondegeneracy at a local minimizer of the initial problem. They use essentially the characterization of strong regularity for such problems given by Bonnans and Ramírez [2], however some "weak" strict complementarity condition has to be supposed.

The paper is structured as follows. Section 2 is devoted to some preliminaries and basic notation. In section 3 we prove that nondegeneracy of a stationary solution and uniqueness of the multiplier for (1) necessarily follow from the Aubin property for the critical point map Σ. This has been known before only for special cases, cf. [5,12, 10]. In section 4 we recall from [10, Ch. 4] the result on single-valuedness under the Aubin property for abstract global minimizing set mappings and apply this to convex cone constrained programs. Finally, section 5 is concerned with the equivalence of strong and metric regularity for critical point systems of (1) in the case of K being described by a finite system of nonlinear inequalities. This extends the well-known results for a convex polyhedral set K, cf. [5,10].

2 Preliminaries

In this section we introduce further notation and terminology of this paper, and we provide some basic tools and results which are needed in the next sections.

Let $C \subset \mathbb{R}^m$ be a convex set. The relative interior of C is denoted by $\mathrm{ri}\, C$, while $\mathrm{span}\, C$ is the linear hull of C. A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is called *convex with respect to C* if the graph of the multifunction $g(x) + C$ is convex.

If C is a convex closed cone, $\mathrm{lin}\, C$ denotes the largest subspace in C, $C^- = \{y^* \in \mathbb{R}^m \mid \langle y^*, y\rangle \le 0 \ \forall y \in C\}$ is the polar of C, and one has, by classical convex analysis, $(C^-)^- = C$ and

$$(\mathrm{lin}\, C)^\perp = \mathrm{span}(C^-)\,, \tag{3}$$

where $^\perp$ refers to the orthogonal complement. Given a linear subspace X of $\mathbb{R}^m$, and writing A^* for the adjoint of a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$, we obviously have

$$A\mathbb{R}^n + X = \mathbb{R}^m \qquad \text{if and only if} \qquad \left\{ \begin{array}{c} A^* u = 0 \\ u \in X^\perp \end{array} \right\} \Rightarrow \ u = 0\,. \tag{4}$$

Given a closed convex set $K \subset \mathbb{R}^m$, then $T_K(\bar{y}) = N_K(\bar{y})^-$ denotes the *tangent cone* of K at $\bar{y}$, where $N_K(\bar{y})$ is the normal cone as above. Again, $T_K(\bar{y})$ is empty by definition if $\bar{y} \notin K$.

If K is a closed convex cone, one has $\lambda \in N_K(y)$ if and only if $y \in N_{K^-}(\lambda)$, i.e., in this case (2) is equivalent to

$$D_x L(\bar{x}, \lambda) = a, \quad g(\bar{x}) - b \in N_{K^-}(\lambda), \tag{5}$$

where obviously $N_{K^-}(\lambda) = K \cap \{\mu \in \mathbb{R}^m \mid \langle \mu, \lambda \rangle = 0\}$.

We will refer in this paper at some places to a basic theorem on the equivalence of metric and strong regularity given in [5]. Consider the perturbed generalized equation

$$0 \in p + F(z,t) + N_C(z), \tag{6}$$

where $C \subset \mathbb{R}^d$ is a nonempty, polyhedral, convex set, $F : \mathbb{R}^d \times \mathbb{R}^\kappa \to \mathbb{R}^d$ is a locally Lipschitz function which is continuously differentiable w.r. to z, and $q = (p,t)$ is a parameter vector varying around some initial point $q^0 = (p^0, t^0)$. Denote the solution set of (6) by $\widetilde{S}(q)$. Given $(q^0, z^0) \in \operatorname{gph} \widetilde{S}$, a linearized version of the generalized equation (6) is given by

$$0 \in \pi + D_z F(z^0, t^0) z + N_C(z), \tag{7}$$

where the canonical parameter π varies around $\pi^0 = p^0 + F(z^0, t^0) - D_z F(z^0, t^0) z^0$. Let $\widetilde{L}(\pi)$ denote the solution set of (7).

Proposition 1 (Dontchev and Rockafellar [5, Thm. 3]) *Given $z^0 \in \widetilde{S}(q^0) = \widetilde{L}(\pi^0)$, the following properties are equivalent:*

1. *the solution map $\widetilde{S}$ of (6) has the Aubin property at (q^0, z^0),*
2. *the solution map $\widetilde{S}$ of (6) is locally single-valued and Lipschitz around (q^0, z^0),*
3. *the solution map $\widetilde{L}$ of the linearization (7) has the Aubin property at (π^0, z^0),*
4. *the solution map $\widetilde{L}$ of the linearization (7) is locally single-valued and Lipschitz around (π^0, z^0).*

Note that property 4 is the concept of *strong regularity in Robinson's [15] sense.* Proposition 1 applies to the KKT system (2) of the cone constrained program (1) if K is a convex polyhedral set: this is obvious if K is in addition a cone, due to (5), otherwise a suitable transformation is needed, for details see [5].

In section 5 we will apply some basic result on persistence of metric or strong regularity of a continuous function $F : \mathbb{R}^d \to \mathbb{R}^s$ with respect to small Lipschitz perturbations $F - g$. Given a solution $\bar{z}$ of the equation $F(z) = 0$, let $\Omega \subset \mathbb{R}^d$ be some neighborhood of $\bar{z}$ and let $g : \Omega \to \mathbb{R}^s$ be Lipschitz on Ω. We put

$$\begin{aligned} \sup(g,\Omega) &= \sup\{\|g(z)\|' \mid z \in \Omega\}, \\ \operatorname{Lip}(g,\Omega) &= \inf\{r > 0 \mid \|g(z) - g(z')\|' \le r\|z - z'\| \ \forall z, z' \in \Omega\}, \end{aligned}$$

where $\|\cdot\|$ and $\|\cdot\|'$ are given norms in $\mathbb{R}^d$ and $\mathbb{R}^s$, respectively. $\operatorname{Lip}(g,\Omega)$ is called the *Lipschitz rank* of g. The space $G = C^{0,1}(\Omega, \mathbb{R}^s)$ of our (locally) Lipschitz variations g is supposed to be equipped with the norm

$$|g| = \max\{\sup(g,\Omega), \operatorname{Lip}(g,\Omega)\}. \tag{8}$$

Persistence of *strong regularity* for generalized equations was originally studied by Robinson [15] for small C^1-functions g. The following proposition was given by Kummer [13] (see also [10, §4.1]) in the more general context of perturbed inclusions, in our special form it already follows from Cominetti [4].

Proposition 2 (Klatte and Kummer [10, Cor. 4.4]) *Let F be a continuous function from $\mathbb{R}^d$ to $\mathbb{R}^s$. Suppose that $g \in G$ satisfies* $\sup(g, \bar{z}+rB) = o(r)$ *and* $\operatorname{Lip}(g, \bar{z}+rB) = O(r)$. *Then F is metrically (resp. strongly) regular regular at $\bar{z}$ if and only if so is $F - g$.*

Here $0(\cdot)$ and $o(\cdot)$ denote as usual functions of the type $0(t) \to 0$ and $o(t)/t \to 0$ for $t \to 0$ with $0(0) = 0$ and $o(0) = 0$, while B is the unit ball in $\mathbb{R}^d$.

3 Constraint nondegeneracy under Aubin property

In this section, we assume that f and g are C^1 functions and that $K \subset \mathbb{R}^m$ is a closed convex set. We shall show the Aubin property of the critical point map Σ implies that the Lagrange multiplier is unique.

According to [1,18] a point $\bar{x}$ satisfying $g(\bar{x}) \in K$ is called *nondegenerate* with respect to g and K if

$$Dg(\bar{x})\mathbb{R}^n + \operatorname{lin} T_K(g(\bar{x})) = \mathbb{R}^m. \tag{9}$$

This means surjectivity of $Dg(\bar{x})$ whenever the cone is pointed.

By (3) and (4), condition (9) is equivalent to

$$[\, Dg(\bar{x})^* u = 0 \;\wedge\; u \in \operatorname{span} N_K(g(\bar{x}))\,] \;\Rightarrow\; u = 0\,, \tag{10}$$

cf. [1, §4.6]. For a convex polyhedral cone K, (9) (and hence (10)) coincides with Robinson's [16] definition of nondegeneracy; in the standard NLP case $K = \mathbb{R}^k_- \times \{0_{m-k}\}$ this is the linear independence constraint qualification (LICQ).

It is well-known (cf. e.g. [18, Thm. 2.1]) that

$$\big[(\bar{x},\bar{\lambda}) \in \Sigma(0) \;\wedge\; \bar{x} \text{ nondegenerate w.r. to } g \text{ and } K\big] \;\;\Rightarrow\;\; \Lambda(\bar{x}\} = \{\bar{\lambda}\}, \tag{11}$$

and conversely, $\Lambda(\bar{x},0) = \{\lambda\}$ and $\lambda \in \operatorname{ri} N_K(g(\bar{x}))$ together imply that $\bar{x}$ is nondegenerate w.r. to g and K.

To see (11) immediately, let $\lambda^1, \lambda^2 \in \Lambda(\bar{x},0)$. Then $Dg(\bar{x})^*\lambda^i = -Df(\bar{x})$ and $\lambda^i \in N_K(g(\bar{x}))$ for $i = 1,2$, hence $Dg(\bar{x})^*(\lambda^1 - \lambda^2) = 0$ and $\lambda^1 - \lambda^2 \in \operatorname{span} N_K(g(\bar{x}))$, which implies $\lambda^1 = \lambda^2$ because of (10) and we are done.

Theorem 1 *Given a critical point $(\bar{x},\bar{\lambda}) \in \Sigma(0)$, suppose Σ has the Aubin property at $(0,(\bar{x},\bar{\lambda}))$. Then $\bar{x}$ is nondegenerate with respect to g and K, and so $\Lambda(\bar{x},0) = \{\bar{\lambda}\}$.*

Proof: Let $(\bar{x},\bar{\lambda}) \in \Sigma(0)$ and $u \in \mathbb{R}^m$ such that

$$Dg(\bar{x})^* u = 0 \;\wedge\; u \in \operatorname{span} N_K(g(\bar{x})).$$

We have to show that $u = 0$. If $N_K(g(\bar{x})) = \{0\}$ there is nothing to prove. Suppose $N_K(g(\bar{x})) \neq \{0\}$ and choose any $w \in \operatorname{ri} N_K(g(\bar{x}))$ with $\|w\| = 1$. Define for any $\varepsilon, \delta > 0$

$$a(\varepsilon) = \varepsilon Dg(\bar{x})^* w, \qquad b(\delta) = \delta u.$$

Then we have $\bar{\lambda} + \varepsilon w \in N_K(g(\bar{x}))$ from $\bar{\lambda}, w \in N_K(g(\bar{x}))$, and $D_x f(\bar{x}) + Dg(\bar{x})^* \bar{\lambda} = 0$ implies $D_x f(\bar{x}) + Dg(\bar{x})^*(\bar{\lambda} + \varepsilon w) = a(\varepsilon)$. Thus,

$$(\bar{x}, \bar{\lambda} + \varepsilon w) \in \Sigma(a(\varepsilon), 0).$$

For small fixed ε, and for all $\delta \downarrow 0$, there are such elements $(x, \lambda) \in \Sigma(a(\varepsilon), b(\delta))$ which satisfy with some Lipschitz constant L the inequality of the Aubin property,

$$\|(x, \lambda) - (\bar{x}, \bar{\lambda} + \varepsilon w)\| \leq L\delta \|u\|. \tag{12}$$

By construction, $\bar{\lambda} + \varepsilon w \in \operatorname{ri} N_K(g(\bar{x}))$ and $u \in \operatorname{span} N_K(g(\bar{x}))$, hence we can choose some small, but fixed $t > 0$ such that

$$\bar{\lambda} + \varepsilon w - tu \in N_K(g(\bar{x})).$$

This implies, together with $g(x) - b(\delta) = g(x) - \delta u \in K$,

$$\begin{aligned} 0 &\geq \langle \bar{\lambda} + \varepsilon w - tu, g(x) - \delta u - g(\bar{x}) \rangle \\ &= \langle \bar{\lambda} + \varepsilon w - tu - \lambda, g(x) - \delta u - g(\bar{x}) \rangle + \langle \lambda, g(x) - \delta u - g(\bar{x}) \rangle \\ &\geq \langle \bar{\lambda} + \varepsilon w - \lambda - tu, g(x) - g(\bar{x}) - \delta u \rangle, \end{aligned}$$

where the last inequality follows from $\lambda \in N_K(g(x) - \delta u)$ and $g(\bar{x}) \in K$. Continuing this, we obtain, with $\mu = \bar{\lambda} + \varepsilon w - \lambda$,

$$t\delta \|u\|^2 \leq \langle \mu, g(\bar{x}) - g(x) \rangle + t\langle u, g(x) - g(\bar{x}) \rangle + \delta \langle \mu, u \rangle. \tag{13}$$

By applying the mean-value theorem to the components g_i of g, one has

$$g_i(x) - g_i(\bar{x}) = Dg_i(\xi^i)(x - \bar{x}) \quad \text{with some } \xi^i \text{ between } x \text{ and } \bar{x}.$$

Now the three terms in the right-hand side sum of (13) can be estimated as follows, recall that $\mu = \bar{\lambda} + \varepsilon w - \lambda$ and $x - \bar{x}$ fulfil the Lipschitz estimate (12). Indeed,

$$\langle \mu, g(\bar{x}) - g(x) \rangle \;=\; \sum_i \mu_i Dg_i(\xi^i)(x - \bar{x}) \;\leq\; \sum_i |\mu_i| \, \|Dg_i(\xi^i)\| \, \|x - \bar{x}\| \;=\; o(\delta)$$

is obtained after applying the Lipschitz estimate (12) twice. Again by (12),

$$\delta \langle \mu, u \rangle \;\leq \delta \, \|\mu\| \, \|u\| \;\leq\; L\delta^2 \|u\|^2 = o(\delta).$$

Finally, we have $Dg_i(\xi^i) \to Dg_i(\bar{x})$, since Dg is continuous and $\xi^i \to \bar{x}$ as $\delta \downarrow 0$, i.e.

$$\langle u, g(x) - g(\bar{x}) \rangle \;=\; t \sum_i u_i Dg_i(\xi^i)(x - \bar{x})$$

with $\sum_i u_i Dg_i(\xi^i)$ tending to $Dg(\bar{x})^* u = 0$ as $\delta \downarrow 0$. This yields

$$t\langle u, g(x) - g(\bar{x}) \rangle \;\leq\; t \, \|\sum_i u_i Dg_i(\xi^i)\| \, \|x - \bar{x}\| \;=\; o(\delta).$$

Thus, (13) says, because of $t > 0$,

$$\delta \|u\|^2 = o(\delta)$$

for arbitrarily small δ. This implies $u = 0$ and we are done. □

Remark 1 (Generalized equations). Theorem 1 similarly holds for the generalized equation

$$F(x)+Dg(x)^*\lambda = a, \quad \lambda \in N_K(g(x)-b), \tag{14}$$

with g, K as above and some continuous function F. Its proof and that of (11) made nowhere use of the special form $F(x) = Df(x)$ in (2). If K is a convex polyhedral set and F and Dg are continuously differentiable, then Theorem 1 is a special consequence of Proposition 1.

Remark 2 (Generalized Kojima systems). If the generalized equation (14) can be reformulated in equation form with a so-called (generalized) Kojima function, then Theorem 1 is Thm. 7.1 in [10], cf. also [12]. Note that this reformulation becomes possible if K itself is described by a system of smooth inequalities (under some constraint qualification for this system), for details see [10, Chapt. 7]. Finally, in the context of nonlinear semidefinite programs, Fusek [8] has shown the uniqueness of the multiplier under the Aubin property of critical points.

4 Aubin property versus uniqueness for global minimizers

In this section, we recall a basic result on uniqueness of global minimizers to abstract optimization problems under the Aubin property and apply this to convex cone constrained programs of the form (1).

For the moment, we consider the abstract perturbed optimization problem

$$\begin{aligned} &\min\ \varphi(x,t) - \langle a,x\rangle \ \text{ s.t. } \ x \in M(t) \subset \mathscr{M} \quad (\varphi : \mathscr{M}\times T \to \mathbb{R}),\\ &\emptyset \neq \mathscr{M} \subset X,\ (X,\langle\cdot,\cdot\rangle) \text{ Hilbert space, } (T,d(\cdot,\cdot)) \text{ metric space,} \end{aligned} \tag{15}$$

where $\bar{t} \in T$ is an initial parameter, and (a,t) varies in $X \times T$ near $(0,\bar{t}) \in X \times T$ measured by $\rho((a,t),(a',t')) = \|a-a'\| + d(t,t')$. Define by

$$\Psi(a,t) = \operatorname{argmin}_x\{\varphi(x,t) - \langle a,x\rangle \,|\, x \in M(t)\}$$

the set of (global) optimal solutions of problem (15), and let $\Phi(a) = \Psi(a,\bar{t})$.

Lemma 1 [10, Lemma 4.6]. *Given $\bar{x} \in \Phi(0)$, suppose that* $\operatorname{dist}(\bar{x},\Phi(a))$ *converges to zero for each sequence $a \to 0$, i.e., Φ is lower semicontinuous at $(0,\bar{x})$. Then $\Phi(0) = \{\bar{x}\}$.*

Inspecting the proof of Lemma 4.6 in [10], one sees that the ”rich” (”tilt”) perturbation $\langle a,x\rangle$ in the objective function is crucial. Then one immediately has the following result which is a modification of Corollary 4.7 in [10].

Theorem 2 (Aubin property and locally single-valued solutions). *In the setting (15), the solution mapping Ψ has the Aubin property at $((0,\bar{t}),\bar{x}) \in \operatorname{gph}\Psi$ only if it is single-valued near $(0,\bar{t})$.*

Proof It follows from the Aubin property at $((0,\bar{t}),\bar{x})$ that for certain neighborhoods U of $(0,\bar{t})$ and V of $\bar{x}$ and for all $(a,t) \in U$ and $x \in V$ with $x \in \Psi(a,t)$,

$$\Psi(a,t) \cap V \neq \emptyset \quad \text{and} \quad \operatorname{dist}(x,\Psi(a',t')) \to 0 \ \text{ as } \ (a',t') \to (a,t).$$

Then Lemma 1 yields $\Psi(a,t) = \{x(a,t)\}$ for all $(a,t) \in U$ and some $x(a,t) \in V$. □

Next we consider the perturbed cone constrained optimization problem (1) and suppose in addition to the general assumptions for (1) that f and g are continuously differentiable, f is a convex function and g is convex with respect to the set $-K$. Let us speak in this case of the *convex problem (1) with C^1 data.*

Hence, by standard convex optimization (cf. e.g. [1, Prop. 3.3]), one has: if some $(\bar{x},\lambda)$ satisfies the KKT conditions (2), then $\bar{x}$ is a global minimizer of the optimization problem $P(p)$. Recall that $\Sigma(p)$ (resp. $S(p)$) is the set of critical points (resp.stationary solutions) of $P(p)$.

Corollary 1 (Single-valued solutions for convex programs). *Consider the convex problem (1) with C^1 data, and let $(0,(\bar{x},\bar{\lambda})) \in \operatorname{gph}\Sigma$. Then the critical point map Σ is single-valued and Lipschitz on a neighborhood of $p=0$ if and only if Σ has the Aubin property at $(0,(\bar{x},\bar{\lambda}))$.*

Proof The only-if direction follows from the definition of the Aubin property. To show the if-direction, we first observe that the stationary point map S has the Aubin property at $(0,\bar{x})$: Since Σ has the Aubin property, so also its component S at $\bar{x}$. Thus, S is single-valued, say $S(p)=\{x(p)\}$, by convexity and Theorem 2.

Further, by definition, the Aubin property of Σ at $(0,(\bar{x},\bar{\lambda}))$ implies that Σ has the Aubin property at each $(p,(x,\lambda)) \in \operatorname{gph}\Sigma$ near $(0,(\bar{x},\bar{\lambda}))$. Hence, by Theorem 1, we have that $p \mapsto \Lambda(x(p),p)$ is single-valued near $p=0$, and so, Σ is also single-valued near $p=0$. □

Remark 3 (Strong regularity). Corollary 1 allows the application of conditions for strong regularity in the convex setting, when looking for the Aubin property of Σ. So, for example, strong regularity for linear semi-definite programs (i.e., a special convex cone constrained problem) was characterized in [3]. Further, characterizations of strong regularity of the KKT system *at local minimizers* were given for nonlinear problems of the type (1) with C^2 data, by combining nondegeneracy of solutions with strong second-order optimality conditions: The case of standard nonlinear programs is classic, cf. e.g. [11,15,5,1,10], for semi-definite programs see e.g. [1,19], for second-order cone programs we refer to [2]. In the convex setting, this gives also characterizations for the Aubin property of the KKT mapping.

5 Aubin property and locally single-valued critical points

Now we assume that the set K in the problem (1),

$$P(p), p=(a,b): \qquad \min_x \; f(x) - \langle a,x\rangle \quad \text{subject to} \quad g(x)-b \in K,$$

is defined by finitely many inequalities,

$$K = \{y \in \mathbb{R}^m \,|\, h_j(y) \le 0,\; j=1,\dots,r\}, \quad h_j : \mathbb{R}^m \to \mathbb{R}, \tag{16}$$

write $h=(h_1,\dots,h_r)$. We suppose throughout that f,g,h are C^2 functions. Convexity of K is not required. Smooth equations could be added, we avoid this for simplicity.

In this section, we will discuss consequences of the Aubin property of the solution map Σ of the KKT system (2) if K is described by (16).

For $y \in K$ and under some constraint qualification for the system $h(y) \le 0$ (e.g., under the Mangasarian-Fromovitz constraint qualification (MFCQ)), the normal cone $N_K(y)$ has the representation, cf. e.g. [10, §7.2],

$$\lambda \in N_K(y) \;\Leftrightarrow\; [\, \exists \mu \in \mathbb{R}^r : \lambda = Dh(y)^* \mu^+ \;\wedge\; h(y) = \mu^- \,], \tag{17}$$

where $\mu_j^+ = \max\{\mu_j, 0\}$ and $\mu_j^- = \min\{\mu_j, 0\}$. This is simply a brief description of the cone of the active gradients $Dh_j(y)$, $j \in I(y)$, with $I(y) = \{j \,|\, h_j(y) = 0\}$.

For the rest of this section, let $\bar{x}$ be a stationary solution of the problem (P)=$P(0)$, and put $\bar{g} = g(\bar{x})$. Suppose that MFCQ is satisfied for the system $h(y) \le 0$ at $y = \bar{g}$. Setting

$$G(x) = h(g(x)), \;\text{ which gives }\; DG(x)^* = Dg(x)^* Dh(g(x))^*, \tag{18}$$

formula (17) allows to write the KKT system (2) for the initial problem (P) at $x = \bar{x}$ and $y = \bar{g}$ in the following equivalent (Kojima [11]) form

$$F(x,\mu) = \begin{pmatrix} Df(x) + DG(x)^* \mu^+ \\ G(x) - \mu^- \end{pmatrix} = 0, \quad \lambda \text{ from (17)}. \tag{19}$$

System (19) is the Kojima system for the optimization problem

$$\min f(x) \text{ such that } G(x) \le 0 \;\in \mathbb{R}^r. \tag{20}$$

With canonical perturbations (a,c), the equation

$$F(x,\mu) = \begin{pmatrix} a \\ c \end{pmatrix} \in \mathbb{R}^{n+r} \tag{21}$$

describes (equivalently) the KKT-points of

$$\min f(x) - \langle a,x \rangle \;\text{ such that } G(x) \le c. \tag{22}$$

For the classical perturbations (22), the Aubin property and strong regularity are equivalent at the reference point by applying Proposition 1 to a standard nonlinear program with C^2 data f,g.

We also know (Thm. 1) that for both properties LICQ (i.e., $\operatorname{rank} DG(\bar{x}) = r$) is a necessary condition, which does not hold, e.g. if $r > n$. However, LICQ is needed as long as *all* parameters $c \in \mathbb{R}^r$ are taken into account which does not happen here. Indeed, we have

$$\begin{array}{lll} g(x) - b \in K \quad \Leftrightarrow \quad h(g(x)-b) \le 0; & & g(x),\ b \in \mathbb{R}^m \\ h:\ \mathbb{R}^m \to \mathbb{R}^r, \quad g:\ \mathbb{R}^n \to \mathbb{R}^m & & G:\ \mathbb{R}^n \to \mathbb{R}^r \end{array} \tag{23}$$

Setting $r(x,b) = G(x) - h(g(x)-b)$, the constraints $h(g(x)-b) \le 0$ or $g(x) - b \in K$ coincide with

$$G(x) \le \hat{c} := r(x,b). \tag{24}$$

Denote by $\ker A$ and $\operatorname{im} A$ the kernel and the image of a linear operator A, respectively. Recall that nondegeneracy (9) required $Dg(\bar{x})\mathbb{R}^n + \operatorname{lin} T_K(\bar{g}) = \mathbb{R}^m$ and that $\operatorname{lin} T_K(\bar{g}) = \ker Dh(\bar{g})$. So (9) means, in the current context,

$$Dg(\bar{x})\mathbb{R}^n + \ker Dh(\bar{g}) = \mathbb{R}^m. \tag{25}$$

This has consequences for uniqueness of the Lagrange multipliers μ^+ in (19) (μ^- as slack variable is always unique).

Lemma 2 *Under nondegeneracy (9), the Lagrange multiplier μ^+ at the reference point $\bar{x}$ is uniquely defined up to the orthogonal space $(im\, Dh(\bar{g}))^\perp$ of $im\, Dh(\bar{g})$. Hence, supposing $\mu^+ \in im\, Dh(\bar{g})$ in (19), μ^+ is unique at the reference point.*

Proof Assume that μ is not unique at $\bar{x}$, i.e.,

$$\lambda = \sum_{j \in J(\bar{x})} \alpha_j DG_j(\bar{x}) = \sum_{j \in J(\bar{x})} \beta_j DG_j(\bar{x}), \quad \alpha \neq \beta,\ \alpha, \beta \geq 0$$

holds for some $\lambda \in N_K(\bar{g})$, where $J(\bar{x}) = \{j \,|\, G_j(\bar{x}) = h_j(g(\bar{x})) = 0\}$. This implies

$$0 = \sum_{j \in J(\bar{x})} (\alpha_j - \beta_j) Dh_j(\bar{g}) Dg(\bar{x}). \tag{26}$$

Using (25) we know: For any $y \in \mathbb{R}^m$ there are $u \in \mathbb{R}^n$ and $w \in \ker Dh(\bar{g})$ with $Dg(\bar{x})u + w = y$. Multiplying in (26) with u, then yields

$$0 = \sum_{j \in J(\bar{x})} (\alpha_j - \beta_j) Dh_j(\bar{g})(y - w) = \sum_{j \in J(\bar{x})} (\alpha_j - \beta_j) Dh_j(\bar{g}) y.$$

This holds for all $y \in \mathbb{R}^m$. Thus (26) implies that $\alpha - \beta$, with zero-components for $j \notin J(\bar{x})$, is necessarily orthogonal to $Dh(\bar{g})\mathbb{R}^m = \mathrm{im}\, Dh(\bar{g})$. Conversely, if $(\alpha - \beta) \perp \mathrm{im}\, Dh(\bar{g})$ then (26) is evident. □

Next we suppose nondegeneracy (9). F can be written in product form as in [10, §7.1], namely, with identity matrix I,

$$F(x,\mu) = \begin{pmatrix} Df(x) & DG(x)^* & 0 \\ G(x) & 0 & -I \end{pmatrix} \begin{pmatrix} 1 \\ \mu^+ \\ \mu^- \end{pmatrix} = \mathscr{M}(x)\, \mathscr{V}(\mu). \tag{27}$$

With the (partial) linearization of F at $\bar{x}$

$$F_L(x,\mu) = [\mathscr{M}(\bar{x}) + D\mathscr{M}(\bar{x})(x - \bar{x})]\, \mathscr{V}(\mu) \tag{28}$$

let us study the parametrization

$$F_L(x,\mu) = \begin{pmatrix} a \\ c \end{pmatrix} \in \mathbb{R}^{n+r} \quad \text{for } c, \mu \in \mathrm{im}\, Dh(\bar{g}) \text{ only.} \tag{29}$$

Notice that

$$c \in \mathrm{im}\, Dh(\bar{g}) \text{ and } DG(\bar{x})(x - \bar{x}) \in \mathrm{im}\, Dh(\bar{g})$$
$$\text{yield} \quad \mu^- = DG(\bar{x})(x - \bar{x}) - c \in \mathrm{im}\, Dh(\bar{g})\,.$$

Hence $\mu = \mu^+ + \mu^- \in \mathrm{im}\, Dh(\bar{g})$ holds exactly if $\mu^+ \in \mathrm{im}\, Dh(\bar{g})$. For the polyhedral system (29), metric and strong regularity at $((\bar{x},\bar{\mu}),(0,0))$ coincide by Proposition 1.

Now it follows from Proposition 2 that metric and strong regularity at $(\bar{x},\bar{\mu})$ also coincide for the (not linearized) system

$$F(x,\mu) = \begin{pmatrix} a \\ c \end{pmatrix} \in \mathbb{R}^{n+r} \quad \text{for } c, \mu \in \mathrm{im}\, Dh(\bar{g})\,. \tag{30}$$

Indeed, by the C^2 assumptions on f, g, h, the function ϕ as

$$\phi(x) := \mathscr{M}(x) - [\mathscr{M}(\bar{x}) + D\mathscr{M}(\bar{x})(x - \bar{x})]$$

is an arbitrary small Lipschitz function of x. This means that for arbitrarily small $\varepsilon > 0$ there is some $\delta(\varepsilon) > 0$ such that ϕ has the Lipschitz-rank $\varepsilon > 0$ on the ball $\{x \,|\, \|x-\bar{x}\| < \delta(\varepsilon)\}$. Obviously, then also $\sup_{\|x-\bar{x}\|<\delta(\varepsilon)} \|\phi(x)\| \leq \varepsilon\delta(\varepsilon)$ is small. The same holds for $F(x,\mu) - F_L(x,\mu)$ by local boundedness of μ (because of MFCQ). By Proposition 2, sufficiently small Lipschitzian perturbations of the functions do not change metric and strong regularity of the equations (or inclusions), respectively.

Finally, we consider the solution map Σ of the original system (2)

$$Df(x) + Dg(x)^*\lambda = a, \quad \lambda \in N_K(g(x) - b), \quad \textit{i.e.},$$

$$\begin{array}{c} Df(x) + Dg(x)^*\lambda = a \\ \lambda = Dh(g(x) - b)\mu^+, \quad h(g(x) - b) = \mu^-, \end{array} \tag{31}$$

and ask for the equivalence of the Aubin property and local single-valuedness and Lipschitz continuity of Σ.

Compared with the linearization (29) where the second line requires

$$h(\bar{g}) + DG(\bar{x})(x - \bar{x}) \; - \mu^- = c \quad (Dh(\bar{g})b)$$

now the second line becomes $h(g(x) - b) - \mu^- = c$. We may write

$$h(g(x) - b) = h(\bar{g}) + Dh(\bar{g}) \; (g(x) - \bar{g} - b) + o_h(g(x) - \bar{g} - b)$$

$$= h(\bar{g}) + Dh(\bar{g}) \; (\; Dg(\bar{x})(x - \bar{x}) + o_g(x - \bar{x}) - b \;) + o_h(g(x) - \bar{g} - b)$$

$$= h(\bar{g}) + Dh(\bar{g})Dg(\bar{x})(x - \bar{x}) \; + Dh(\bar{g})(o_g(x - \bar{x}) - b) + o_h(\; Dg(\bar{x})(x - \bar{x}) + o_g(x - \bar{x}) - b \;)$$

$$= h(\bar{g}) + DG(\bar{x})(x - \bar{x}) \; + Dh(\bar{g})(o_g(x - \bar{x}) - b) + o_h(\; Dg(\bar{x})(x - \bar{x}) + o_g(x - \bar{x}) - b \;)$$

$$= \; h(\bar{g}) + DG(\bar{x})(x - \bar{x}) \; + o_{Hg}(x - \bar{x}) + o_{hG}(x - \bar{x} - b \;) - Dh(\bar{g})b,$$

with o-type functions o_h, o_g, o_{Hg} and o_{hG}. Given any $\varepsilon > 0$ we thus obtain by standard mean-value estimates (see e.g. [10, §4.3]) that there is some $\delta > 0$ such that, if $x, x' \in B(\bar{x}, \delta)$ and $\|b\| < \delta$, it holds

$$\|o_{Hg}(x - \bar{x}) - o_{Hg}(x' - \bar{x})\| \leq \varepsilon\|(x - x')\|$$

$$\|o_{hG}(x - \bar{x} - b \;) - o_{hG}(x' - \bar{x} - b \;)\| \leq \varepsilon\|(x - x')\|$$

$$\text{and} \quad \|o_{hG}(x - \bar{x} - b \;)\| \leq \varepsilon\|Dh(\bar{g})b\|.$$

So

$$h(g(x) - b) \;\; \text{and} \;\; h(\bar{g}) + DG(\bar{x})(x - \bar{x}) - Dh(\bar{g})b$$

differ again by an arbitrarily small Lipschitz function (of x and b) only.

The linearized problem (29) requires, in the second line,

$$h(\bar{g}) + DG(x - \bar{x}) \; + \mu^- = c := Dh(\bar{g})b.$$

Notice that $c \neq 0$ if $b \neq 0$. Analogously, the first lines can be estimated (due to $f, g, h \in C^2$). They differ again by an arbitrarily small Lipschitz function only. So, again by Proposition 2, the systems (30) and (31) are strongly and metrically regular at the same time.

Thus, we arrived at the equivalence of the Aubin property and strong regularity for the KKT mapping Σ of the cone constrained program (1) if K is given in the form (16) and f,g,h are C^2 functions. We essentially used that nondegeneracy of the constraints is implied by the Aubin property for Σ (Thm. 1 above), we applied known results on equivalence and persistence of strong and metric regularity (Prop. 1, Prop. 2), and we supposed a suitable constraint qualification for the system $h(y) \le 0$ in order to handle the normal cone map $N_K(\cdot)$.

References

1. J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.
2. J.F. Bonnans ans H. Ramírez. Perturbation analysis of second-order cone programming problems, *Math. Program. Ser. B*. 104: 205-227, 2005.
3. Z.X. Chan and D. Sun. Constraint nondegeneracy, strong regularity and nonsingularity in semidefinite programming, *SIAM J. Optim.*. 19: 370-396, 2008.
4. R. Cominetti. Metric regularity, tangent sets and second-order optimality conditions. *Applied Mathematics and Optimization*, 21:265–287, 1990.
5. A. Dontchev and R.T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.*, 6:1087–1105, 1996.
6. A. Dontchev and R.T. Rockafellar. *Implicit Functions and Solution Mappings*. Springer, New York, 2009.
7. P. Fusek. Isolated zeros of Lipschitzian metrically regular R^n functions. *Optimization*, 49:425–446, 2001.
8. P. Fusek. Personal Communication, December 2008.
9. J.-B. Hiriart-Urruty and C. Lemaréchal. Convex Analysis and Minimization Algorithms I,II. Springer, Berlin-Heidelberg, 1993.
10. D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization*. Kluwer Academic Publ., Dordrecht-Boston-London, 2002.
11. M. Kojima. Strongly stable stationary solutions in nonlinear programs. In S.M. Robinson, ed., *Analysis and Computation of Fixed Points*, 93–138. Academic Press, New York, 1980.
12. 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.
13. B. Kummer. Metric regularity: Characterizations, nonsmooth variations and successive approximation. *Optimization*, 46:247–281, 1999.
14. J.V. Outrata and H. Ramírez. The Aubin property of critical points to perturbed second-order cone programs, *SIAM J. Optim.*. 21: 798-823, 2011.
15. S.M. Robinson. Strongly regular generalized equations. *Math. Oper. Res.*, 5:43–62, 1980.
16. S.M. Robinson. Local structure of feasible sets in nonlinear programming: II: Nondegeneracy, *Math. Progr. Study*, 24: 217-230, 1984.
17. R.T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.
18. A. Shapiro. Sensitivity analysis of generalized equations, J. Math. Sci., 115: 2554-2565, 2003.
19. D. Sun. The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications, *Math. Oper. Res.*. 31: 761-776, 2006.