Journal of Convex Analysis
Volume 23 (2016), No. 4, 1247–1262

A Condition Number Theorem in Convex Programming without Uniqueness

Tullio Zolezzi
Dipartimento di Matematica, Università di Genova,
via Dodecaneso 35, 16146 Genova, Italy
zolezzi@dima.unige.it

Dedicated to the memory of Jean Jacques Moreau.

Received: March 24, 2015
Revised manuscript received: December 21, 2015
Accepted: December 29, 2015

A condition number of mathematical programming problems with convex data is defined as a suitable measure of the sensitivity of their optimal solutions under canonical perturbations. A pseudo-distance among mathematical programming problems is introduced via the corresponding Kojima functions. Characterizations of well-conditioning are obtained. We prove that the distance to ill-conditioning is bounded from above by a multiple of the reciprocal of the condition number, thereby generalizing previous results dealing with problems with a unique optimal solution.

Keywords: Convex programming, condition number, distance to ill-conditioning, condition number theorem

2010 Mathematics Subject Classification: 90C31, 90C25

1. Introduction

Consider the finite-dimensional convex programming problem Q, to minimize the objective function $f(x)$ subject to the inequality constraints

$$g_1(x) \leq 0, ..., g_q(x) \leq 0.$$

The purpose of this paper is to generalize in part the condition number theorem obtained in [19] to the case when the considered problems have many optimal solutions. We define the (absolute) condition number of problem Q as a measure of the sensitivity of its optimal solutions with respect to a given class of data perturbations. While the choice of data perturbations is in principle quite arbitrary, we consider here (as in [19]) the canonically perturbed problems $Q(p)$ associated to Q, defined by the parameter $p = (a, b)$, with the objective function $f(x) - a^T x$ and constraints

$$g_1(x) \leq b_1, ..., g_q(x) \leq b_q$$

ISSN 0944-6532 / \$ 2.50

where $(b_1, ..., b_q)^T = b$. Such canonical perturbations play a significant role in the analysis of mathematical programming problems, see [8, Chapter 8]. For a given canonical perturbation p we consider the set $M(p)$ of the minimizers of $Q(p)$, the set $U(p)$ of the corresponding multipliers and the set $S(p) = M(p) \times U(p)$ of the optimal solutions of $Q(p)$. Then the condition number of Q is defined as the Lipschitz modulus of S at $p = 0$ with respect to the Hausdorff distance. Problem Q is then declared well-conditioned if its condition number is finite, ill-conditioned otherwise.

The condition number of some optimization problems (and of other mathematical problems as well) is known to be related to a suitably defined distance of Q from ill-conditioning, starting with the classical Eckart-Young theorem in its optimization form (see [16]). The same property is known to be true in unconstrained infinite-dimensional problems (see [17]), in some class of infinite-dimensional linear-quadratic optimization problems (see [18]), and in convex programming problems with uniqueness (see [19]). The link between condition number and distance to ill-conditioning is given by a lower and an upper bound of the distance in terms of the reciprocal of the condition number, up to suitable constants depending of problem's data. Such a pair of inequalities is called a condition number theorem. The (pseudo-)distance between two programs is defined here by the Lipschitz constant of the difference between their Kojima functions (see [8]), as in [19].

Such results are of interest for the analysis of computational complexity (see [4] and [5]), for the sensitivity and stability properties of the considered problems, and for evaluating the performance of numerical methods. In [18] it is shown however that there exist minimum norm least squares problems for which a condition number theorem (in the above form) cannot hold. This adds interest in finding classes of optimization problems for which a condition number theorem is available. To the knowledge of the author, no such result is known for convex programming problems beyond [19]. We remark that the results of Renegar [12] and [13] deal with the distance to instability (a different notion than ill-conditioning) within linear programming problems (following a different approach). A geometric version of Renegar's condition number is presented in [1]. In [2] a condition number theorem is obtained in the infinite-dimensional setting for unconstrained problems with respect to general perturbations. Extensions to multiobjective optimization are dealt with in [3].

Aim of this paper is to extend the upper estimate of the condition number theorem of [19] to problems with many optimal solutions. We characterize well-conditioning by Lipschitz continuity of the set-valued mapping inverse of the Kojima function. A further characterization is obtained via its coderivative. Upper bounds of the distance to ill-conditioning and of the condition number are obtained through constants related to problem's data, making use of the graphical derivative of the Kojima function. A partial condition number follows.

In Section 2 we collect the relevant notations, assumptions and definitions. Section 3 presents some examples of well- and ill-conditioned problems. An inaccuracy contained in [19] is corrected. Section 4 is devoted to the estimates and the main results. In Section 5 we collect auxiliary results and the proofs. Due to lack of uniqueness, some proofs follow a different pattern than in [19], requiring some results from set-valued analysis and generalized differentiation.

2. Notations, problem statement and assumptions

We consider functions

$$f, g_1, ..., g_q : R^n \to R$$

and the mathematical programming problem Q, to minimize $f(x)$ subject to the constraints

$$g_1(x) \le 0, ..., g_q(x) \le 0.$$

The dimensions n, q are fixed. To each problem Q it is associated a parameter set $\Omega \subset R^{n+q}$ containing 0 as a cluster point. Let us write for simplicity $p = (a, b) \in R^{n+q}$ instead of $p = (a^T, b^T)^T$, where ^{T} denotes transpose. For every $p = (a, b) \in \Omega, a \in R^n, b \in R^q$ we consider the mathematical programming problem $Q(p)$ of minimizing $f(x) - a^T x$ subject to

$$g_1(x) \le b_1, ..., g_q(x) \le b_q \tag{1}$$

with $(b_1, ...b_q)^T = b$. The choice of Ω is in principle arbitrary; in particular we obtain tilt perturbations of Q by choosing

$$\Omega = \{(a, 0) : a \in R^n\}$$

and right-hand side perturbations of the constraints as in (1) by choosing

$$\Omega = \{(0, b) : b \in R^q\}.$$

Denote by g the vector of components $g_1, ..., g_q$. We write $Q = Q(0) = (f, g)$. The following assumption (A1) will hold throughout the paper.

(A1) $f, g_1, ..., g_q$ are convex and continuously differentiable with Lipschitz continuous gradient on every bounded set of R^n.

Given $p \in \Omega$ we denote by $M(p)$ the set of all minimizers (if any) of $Q(p)$, by $U(p)$ the set of all corresponding (Kuhn-Tucker) multipliers, and by

$$S(p) = M(p) \times U(p) \tag{2}$$

the set of all optimal solutions of $Q(p)$. We consider the norm on R^m given by

$$|x| = \sum_{j=1}^{m} |x_j|.$$

Sometimes, given $Q = (f, g)$ with parameter set Ω, we shall require that

(A2) there exists $\delta > 0$ such that $Q(p)$ has optimal solutions for every $p \in \Omega$ with $|p| \leq \delta$;

(A3) there exists $x_0 \in R^n$ such that

$$g_j(x_0) < 0, \quad j = 1, ..., q; \tag{3}$$

(A4) there exists a constant $L > 0$ such that

$$|g(x'') - g(x')| \leq L|x'' - x'|$$

for every $x', x'' \in M_0 = \cup\{M(p) : p \in \Omega, |p| \leq \delta\}$.

By (A1), (A4) is fulfilled if M_0 is bounded. As well-known, by (A1) the set of all multipliers (if any) corresponding to a given minimizer of $Q(p)$ does not depend on it, see [7, Proposition 3.1.1 p. 317], thus (2) makes sense. If $A \subset R^m$ and $x \in R^m$ we write

$$\text{dist}(x, A) = \inf\{|x - y| : y \in A\},$$

and if $B \subset R^m$ we consider

$$e(A, B) = \sup\{\text{dist}(a, B) : a \in A\}, \qquad \text{hausd}(A, B) = \max\{e(A, B), e(B, A)\}.$$

Given Q, the *condition number* of problem Q is defined by

$$\text{cond}\, Q = \limsup_{p', p'' \to 0} \frac{\text{hausd}[S(p''), S(p')]}{|p'' - p'|} \tag{4}$$

with S given by (2); of course $p', p'' \in \Omega$ in (4). Problem Q will be called *well-conditioned* if (A2) holds and $\text{cond}\, Q < +\infty$. Given $Q = (f, g)$ we consider the Kojima function K of Q (see [8, Chapter 7]), namely $K : R^{n+q} \to R^{n+q}$ given by

$$K(x, y) = \left(\nabla f(x) + \sum_{j=1}^{q} y_j^+ \nabla g_j(x), g_1(x) - y_1^-, ...g_q(x) - y_q^-\right) \tag{5}$$

where $y^+ = \max\{0, y\}, y^- = \min\{0, y\}$, and (5) means that the first n components of $K(x, y)$ are those of $\nabla f(x) + \sum_{j=1}^{q} y_j^+ \nabla g_j(x)$, and the remaining q are given by $g_j(x) - y_j^-, j = 1, ..., q$. Consider now two problems Q_1, Q_2 with corresponding parameter sets Ω_1, Ω_2 and Kojima functions K_1, K_2. Let $K = K_1 - K_2$ and consider

$$D = \{z \in R^{n+q} : K_1(z) \in \Omega_1 \text{ or } K_2(z) \in \Omega_2\}.$$

Then the (pseudo-) distance between Q_1 and Q_2 is defined by

$$\text{dist}(Q_1, Q_2) = \sup\left\{\frac{|K(z'') - K(z')|}{|z'' - z'|} : z', z'' \in D\right\}. \tag{6}$$

We remark that D is nonempty if e.g. Q_1 fulfills the Slater condition (A3) and $Q_1(p)$ has optimal solutions for some $p \in \Omega_1$, as we see from Proposition 5.1.

Definitions (4) and (6) extend those considered in [19] for problems with unique optimal solutions, and in [16] for unconstrained problems (in the finite-dimensional setting). In [19] problems were considered under boundedness conditions we do not require here. Moreover, the distance defined by (6) is a modification of the one in [19] taking into account the different setting here, and can take also the value $+\infty$.

Given Q, if Ω is a neighborhood of 0 in R^{n+q}, so that the full amount of all canonical perturbations is used, and $Q(p)$ has many optimal solutions, then well-conditioning of Q is a rather restrictive property. Indeed, in this case the solution mapping S is Lipschitz continuous around 0 by (4), hence by [9, Corollary 1] S is single-valued around 0. A notion of conditioning based on a less stringent behavior of the optimal solution sets (like upper Lipschitz properties) will be considered elsewhere.

3. Examples

Example 3.1. Let $n = q = 2$ and consider

$$f(x) = x_2, \qquad g_1(x) = x_1^2 + x_2^2 - 1, \qquad g_2(x) = -x_2,$$

$$\Omega = \{(a, b) \in R^4 : a_1 = 0, b_1 \geq 0, b_2 \leq 0, |a| + |b| < 1\}.$$

Then, given $p = (a, b) \in \Omega, Q(p)$ consists of minimizing $(1 - a_2)x_2$ subject to

$$x_1^2 + x_2^2 \leq 1 + b_1, \qquad x_2 \geq -b_2.$$

Thus $M(a, b)$ is the segment $|x_1| \leq \sqrt{1 + b_1 - b_2^2}, x_2 = -b_2$ and $U(a, b)$ is the singleton $(0, 1 - a_2)^T$. Then M, U are Lipschitz continuous near 0 and Q is well-conditioned (by Lemma 5.2).

Example 3.2. This example is borrowed from [8, Example 8.22]. Let $n = 2$, $q = 3$ and consider the problem Q of minimizing x_2 subject to the constraints

$$x_1^2 - x_2 \leq 0, \qquad 0 \leq x_2 \leq 1.$$

If we consider $p = (0, 0, 0, b_2, 0)^T$, then the corresponding set of minimizers of $Q(p)$ is the singleton $0 \in R^2$ if $b_2 \geq 0$ and the segment

$$|x_1| \leq \sqrt{-b_2}, \quad x_2 = -b_2 \text{ if } -1 < b_2 < 0.$$

Hence if $-1 < b_2 < 0$ we have $e[M(p), M(0)] = -b_2 + \sqrt{-b_2}$, whence ill-conditioning with

$$\Omega = \{(a, b) \in R^5 : a = 0, b_i = 0 \text{ if } i \neq 2, b_2 > -1\}.$$

Example 3.3. Let $n = 2$, $q = 1$ and Q be the linear program, to minimize x_2 with the constraint $x_2 \geq 0$. For any $p = (a_1, a_2, b)^T$ with $a_2 < 1, Q(p)$ consists of minimizing $(1-a_2)x_2-a_1x_1$ subject to the constraint $x_2 \geq -b$. Then $S(a,b)=\emptyset$ if $a_1 \neq 0$, and if $a_1 = 0$ we have

$$S(a,b)=\{(x_1,-b)^T : x_1 \in R\} \times \{1-a_2\}.$$

Hence S is Lipschitz on $\Omega = \{(0,a_2,b)^T : a_2 < 1, b \in R\}$, and Q is well-conditioned with such a choice of Ω.

Remark 3.4. Let Q be a linear programming problem with only inequality constraints, and consider perturbing their right-hand side, so that $\Omega = \{0\} \times R^q, 0 \in R^n$. Suppose that the optimal value of $Q(0,b)$ is finite for every $b \in R^q$ sufficiently small. Then the set-valued mapping M is Lipschitz continuous, see [6, p. 153]. By duality (see [7, Chapter VII]) $U(0,b)$ is the set of the optimal solutions of another linear program, however Lipschitz continuity of U with respect to b fails in general (see [10, Remark 2.7]).

Remark 3.5. The statement in Example 1 of [19] is incorrect since [14, Theorem 1] proves only upper Lipschitz behavior of the optimal solutions (under more general perturbations).

Remark 3.6. Suppose that $Q(p)$ has a unique optimal solution for each $p \in \Omega$ sufficiently small. Let Q be well-conditioned. Then inequality (6) in [19] needs to be replaced by

$$\operatorname{dist}(Q,\bar{Q})[(1+L)\operatorname{cond} Q+1]<1$$

so that the correct lower bound in [19, (7)] becomes

$$\operatorname{dist}(Q,IC) \geq \frac{1}{1+(1+L)\operatorname{cond} Q}.$$

The reason is that the term $|b'-b''|$ was omitted in the inequality after (17) of [19], hence in the proof of Theorem 2 there we have $K^* \leq (1+L)\operatorname{cond} Q+1$.

4. Main results

A characterization of well-conditioned mathematical programming problems in our setting can be obtained making use of Lipschitz properties of their Kojima functions. Given the single-valued function K we consider the set-valued mapping K^{-1} inverse of K. For $\delta > 0$ consider

$$\Omega(\delta)=\{p \in \Omega : |p| \leq \delta\}.$$

Theorem 4.1. *Let Q with parameter set Ω fulfill* (A3) *and* (A4). *Then the following properties are equivalent:*

$$Q \textit{ is well-conditioned}\,; \tag{7}$$

$$K^{-1} \textit{ is Lipschitz continuous on } \Omega(\delta) \textit{ for some } \delta > 0. \tag{8}$$

A further characterization of well-conditioning is based on the coderivative of the Kojima function, as follows. For a given set-valued mapping F and points x, y with $y \in F(x)$ we denote by $D^*F(x, y)$ the coderivative of F at (x, y), see [15, Definition 8.33]. Since the Kojima function K is single-valued, we write $D^*K(z)$ if $K(z) = 0$ instead of $D^*K(z, 0)$. A *lower dimensional ball* Σ in R^{n+q} is any set of points $z \in R^{n+q}$ such that $|z| \le \delta$ for some $\delta > 0$ and $z = (z_1, ...z_{n+q})^T \in \Sigma$ iff $z_j = 0, j \in I$ an index set properly contained in $\{1, ..., n + q\}$.

Example 4.2. Consider Example 3.3 and let now

$$\Omega^* = \{(0, a_2, b)^T \in R^3 : |a_2| + |b| < 1\}.$$

Then all conclusions about this example are unchanged with the new choice of the parameter set Ω^*, which is a lower dimensional ball in R^3.

As a consequence of Theorem 4.1 we have the following characterization of well-conditioned problems.

Theorem 4.3. *Let Q fulfill* (A3) *with $\Omega(\delta)$ a lower dimensional ball. Suppose that $\cup\{S(p) : p \in \Omega(\delta)\}$ is bounded. Then the following are equivalent:*

$$Q \text{ is well-conditioned;} \tag{7}$$

$$0 \in D^*K(z)(w) \text{ implies } w = 0 \text{ for every } z \text{ with } K(z) = 0. \tag{9}$$

In order to obtain a partial condition number theorem, we start by establishing an upper bound of the distance of a given problem to ill-conditioning. Denote by IC the set of *ill-conditioned* problems Q^*, namely those which verify (A2), such that cond $Q^* = +\infty$. Write

$$\operatorname{dist}(Q, IC) = \inf\{\operatorname{dist}(Q, Q^*) : Q^* \in IC\}.$$

If $Q = (f, g)$ fulfills (A3), in order to estimate the distance to ill-conditioning we shall consider $\varepsilon > 0$ such that

$$g_j(x_0) \le -\varepsilon, \quad j = 1, ..., q. \tag{10}$$

For simpler statement, in the next Proposition 4.4 we refer to formula (36) at the end of its proof.

Proposition 4.4. *Let $Q = (f, g)$ be with parameter set $\Omega = \Omega(\delta)$, $\delta < \min\{\varepsilon, 1\}$. Suppose that* (A2) *and* (A3) *hold, and that*

$$\sup\{|x| : x \in M(p), p \in \Omega\} = s < +\infty. \tag{11}$$

Then dist $(Q, IC) \le c_1$ where c_1 is given by (36).

Now we derive an upper bound of the condition number. As shown by Lemma 5.3, suffices to estimate the Lipschitz modulus at $p = 0$ of the set-valued mapping inverse of the Kojima function. To this aim, given $Q = (f, g)$ with sufficiently smooth data and Kojima function K, for every $x \in R^n$ and $y \in R^q$ we consider the $q \times n$ Jacobian matrix $\nabla g(x)$, and the $(n+q) \times (n+2q)$ matrix

$$A(x,y) = \begin{pmatrix} \nabla^2 f(x) + \sum_{j=1}^q y_j^+ \nabla^2 g_j(x) & \nabla g(x)^T & 0 \\ \nabla g(x) & 0 & I \end{pmatrix} \tag{12}$$

where ∇^2 denotes the Hessian matrix, the north-east 0 is $n \times q$, the other 0 is $q \times q$ as well as the identity matrix I. Given y and w_2 in R^q let $\alpha \in R^q$ be with the $i-$ th component α_i given by

$$0 \text{ if } y_i < 0, \qquad w_{2i}^+ \text{ if } y_i = 0, \qquad w_{2i} \text{ if } y_i > 0, \tag{13}$$

and denote by $\beta \in R^q$ the vector whose $i-$ th component β_i is given by

$$-w_{2i} \text{ if } y_i < 0, \qquad -w_{2i}^- \text{ if } y_i = 0, \qquad 0 \text{ if } y_i > 0. \tag{14}$$

Consider any $z' = (x', y')$ in $R^{n+q}, w = (w_1, w_2)$ with $w_1 \in R^n, w_2 \in R^q$ and $v \in R^{n+q}$, and define

$$K_*(z') = \sup_{|v| \le 1} \inf\{|w| : v = A(x', y')(w_1, \alpha, \beta)\} \tag{15}$$

where A is given by (12). Finally let c_2 be any real constant such that

$$c_2 > \sup\{\limsup_{z' \to z} K_*(z') : K(z) = 0\}. \tag{16}$$

Proposition 4.5. *Let $Q = (f, g)$ be well-conditioned with $f, g \in C^2(R^n)$ and parameter set a lower dimensional ball Ω. Suppose that* (A3) *holds and that*

$$\cup\{S(p) : p \in \Omega\} \quad \textit{is bounded.} \tag{17}$$

If c_2 verifies (16) *then*

$$\operatorname{cond} Q \le (L+2)c_2 + 1, \tag{18}$$

L as in (A4).

In the following illustrative example, (18) holds even if (17) doesn't.

Example 4.6. Consider again Example 3.3. We have $n = 2, q = 1$ and

$$K(x_1, x_2, y) = (0, 1 - y^+, -x_2 - y^-)^T,$$

$$A(x_1, x_2, y) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$$

Thus, writing $v = (v_1, v_2)$ with $v_1 \in R^2, v_2 \in R$, we have $v = A(x, y)(w_1, \alpha, \beta)$ iff

$$v_{11} = 0, \qquad v_{12} = -\alpha, \qquad v_2 = -w_{12} + \beta.$$

Following (13) and (14) we have three cases. Given $z' = (x', y')$, if $y' > 0$ we compute

$$\inf\{|w_{11}| + |w_{12}| + |w_2| : v_{11} = 0,\ v_{12} = -w_2, v_2 = -w_{12}\} = |v_2| + |v_{12}|.$$

If $y' = 0$ we obtain

$$\inf\{|w_{11}| + |w_{12}| + |w_2| : v_{11} = 0,\ v_{12} = -w_2^+, v_2 = -w_{12} - w_2^-\} \le |v_2| + |v_{12}|.$$

If $y' < 0$

$$\inf\{|w_{11}| + |w_{12}| + |w_2| : v_{11} = 0,\ v_{12} = 0, v_2 = -w_{12} - w_2\} \le |v_2|.$$

It follows that $K_*(z') \le 1$ for every z', hence any $c_2 > 1$ works in (16). We have $\operatorname{cond} Q \le 1$, so that (18) is fulfilled.

From Propositions 4.4 and 4.5 we obtain

Corollary 4.7. *Let Q fulfill the assumptions of Propositions* 4.4 *and* 4.5. *If cond $Q > 0$, then*

$$\operatorname{dist}(Q, IC) \le \frac{c_1[(L+2)c_2 + 1]}{\operatorname{cond} Q} \tag{19}$$

where c_1, c_2 are given by (36), (16) *respectively.*

In a sense, (19) can be considered as one half of a condition number theorem. We have also an estimate of the condition number, similar to (18), making direct use of the coderivative D^* as follows. Consider $v, w \in R^{n+q}$ and write

$$T = \sup_{K(x,y)=0} \sup\{|v| : w \in E_1^T v - D^* E_2(-v), |w| \le 1\} \tag{20}$$

where

$$E_1 = \begin{pmatrix} \nabla^2 f(x) & 0 \\ \nabla g(x) & 0 \end{pmatrix}$$

with the north-east 0 of dimension $n \times q$ and the other $q \times q$, and

$$E_2 = \left(\sum_{j=1}^{q} y_j^+ \nabla g_j(x), -y^- \right).$$

Proposition 4.8. *Let $Q = (f, g), f \in C^2(R^n)$, fulfill* (A2), (A3) *and* (17) *with parameter set a lower dimensional ball. Then Q is well-conditioned if $T < +\infty$, and*

$$\operatorname{cond} Q \le (L+2)c_3 + 1$$

for any real constant $c_3 > T, T$ given by (20).

A version of a partial condition number theorem, similar to Corollary 4.7 can be obtained combining Propositions 4.4 and 4.8. Under the corresponding assumptions about problem Q we get (19) with c_3 replacing c_2.

5. Auxiliary results and proofs

We need some preliminary results, as follows.

Proposition 5.1. *Given problem $Q = (f, g)$ with corresponding Kojima fuction K, let $p = (a, b) \in R^{n+q}$ and suppose that* (A3) *holds. Then the following are equivalent:*

$$x \in M(p) \quad \textit{and} \quad u \in U(p); \tag{21}$$

$$K(x, u + g(x) - b) = p. \tag{22}$$

Proof of Proposition 5.1. As well known, by (A1) and (A3), (21) is equivalent to the optimality system

$$\nabla f(x) + \sum_{j=1}^{q} u_j \nabla g_j(x) = a; \quad u_j \geq 0,$$

$$u_j(g_j(x) - b_j) = 0, \quad g_j(x) \leq b_j, \quad j = 1, ..., q.$$

Then the conclusion follows as in [19, Lemma 1] (where uniqueness of the optimal solution plays no role).

Lemma 5.2. *If the set-valued mappings M, U are Lipschitz continuous on $\Omega(\delta)$ for some $\delta > 0$, then the same is true for $M \times U$.*

Proof of Lemma 5.2. Given $p', p'' \in \Omega(\delta)$ and given $x'' \in M(p''), u'' \in U(p'')$, we have

$$\begin{aligned} &\operatorname{dist}[(x'', u''), M(p') \times U(p')] \\ = \; &\inf\{|x'' - x'| + |u'' - u'| : x' \in M(p'), u' \in U(p')\}. \end{aligned} \tag{23}$$

For every $\varepsilon > 0$ there exist $\bar{x}' \in M(p'), \bar{u}' \in U(p')$ such that

$$\operatorname{dist}[x'', M(p')] + \varepsilon \geq |\bar{x}' - x''|, \qquad \operatorname{dist}[u'', U(p')] + \varepsilon \geq |\bar{u}' - u''|,$$

then by (23)

$$\begin{aligned} &\operatorname{dist}[(x'', u''), M(p') \times U(p')] \leq |\bar{x}' - x''| + |\bar{u}' - u''| \\ \leq \; &2\varepsilon + \operatorname{dist}[x'', M(p')] + \operatorname{dist}[u'', U(p')] \\ \leq \; &2\varepsilon + e[M(p''), M(p')] + e[U(p''), U(p')]. \end{aligned} \tag{24}$$

By taking the supremum in (24) with respect to (x'', u'') we get

$$\begin{aligned} &e[M(p'') \times U(p''), M(p') \times U(p')] \\ \leq \; &\operatorname{hausd}[M(p''), M(p')] + \operatorname{hausd}[U(p''), U(p')] \end{aligned} \tag{25}$$

and similarly for $e[M(p') \times U(p'), M(p'') \times U(p'')]$, hence the conclusion.

Proof of Theorem 4.1. Let (7) hold. Then by (2) and (4) we can assume without restriction that the set-valued mapping $M \times U$ is Lipschitz continuous on $\Omega(\delta), \delta$ as in (A2) (by taking a smaller δ if necessary). By (22) of Proposition 5.1, for every $p = (a, b) \in \Omega(\delta)$

$$K^{-1}(p) = \{(x, u + g(x) - b) : x \in M(p), u \in U(p)\}. \tag{26}$$

Given $p', p'' \in \Omega(\delta)$, $(x'', y'') \in K^{-1}(p'')$, if $(x', y') \in K^{-1}(p')$ we have by (26)

$$y' = u' + g(x') - b', \qquad y'' = u'' + g(x'') - b''$$

with $x' \in M(p')$, $x'' \in M(p'')$, $u' \in U(p')$, $u'' \in U(p'')$. Then remembering (A4)

$$\begin{aligned}
&\operatorname{dist}[(x'', y''), K^{-1}(p')] \\
= \ &\inf\{|x' - x''| + |u'' - u' + g(x'') - g(x') + b' - b''| : x' \in M(p'), u' \in U(p')\} \\
\le \ &(L+1)\inf\{|x' - x''| + |u' - u''| : x' \in M(p'), u' \in U(p')\} + |b' - b''| \\
= \ &(L+1)\operatorname{dist}[(x'', u''), M(p') \times U(p')] + |b' - b''| \\
\le \ &(L+1)\ e[M(p'') \times U(p''), M(p') \times U(p')] + |b' - b''|.
\end{aligned}$$

It follows that

$$e[K^{-1}(p''), K^{-1}(p')] \le (L+1)\ e[M(p'') \times U(p''), M(p') \times U(p')] + |b' - b''|$$

and the analogous inequality exchanging p' with p''. Hence

$$\begin{aligned}
&\operatorname{hausd}[K^{-1}(p'), K^{-1}(p'')] \\
\le \ &(L+1)\operatorname{hausd}[M(p') \times U(p'), M(p'') \times U(p'')] + |b' - b''|,
\end{aligned} \tag{27}$$

thus (27) yields (8) remembering (2). Conversely, assume (8). Lipschitz continuity on the nonempty set $\Omega(\delta)$ implies that $K^{-1}(p) \neq \emptyset$ for each $p \in \Omega(\delta)$, whence (A2) by Proposition 5.1. By assumption there exists a constant $N > 0$ such that

$$e[K^{-1}(p'), K^{-1}(p'')] + e[K^{-1}(p''), K^{-1}(p')] \le N|p' - p''| \tag{28}$$

if $p', p'' \in \Omega(\delta)$. Given any $(x'', y'') \in K^{-1}(p'')$ we have by (26)

$$\inf\{|x' - x''| + |y' - y''| : (x', y') \in K^{-1}(p')\} \ge \inf\{|x' - x''| : x' \in M(p')\}$$

hence

$$\begin{aligned}
&e[K^{-1}(p''), K^{-1}(p')] \\
\ge \ &\sup\{\operatorname{dist}[x'', M(p')] : x'' \in M(p'')\} = e[M(p''), M(p')]
\end{aligned} \tag{29}$$

and similarly for $e[K^{-1}(p'), K^{-1}(p'')]$. Then by (28) we get Lipschitz continuity of M on $\Omega(\delta)$. Now fix any $x'' \in M(p''), u'' \in U(p'')$ and consider $y'' = u'' +$

$g(x'') - b''$. Then $(x'', y'') \in K^{-1}(p'')$ by (26). For every $\varepsilon > 0$ there exists $(x', y') \in K^{-1}(p')$ such that

$$\operatorname{dist}[(x'', y''), K^{-1}(p')] + \varepsilon \geq |x' - x''| + |y' - y''|.$$

We have $u' = y' - g(x') + b' \in U(p')$ and by (A4)

$$|u' - u''| \leq (L + 1)(|x' - x''| + |y' - y''|) + |b' - b''|.$$

It follows that

$$\begin{aligned} \operatorname{dist}[u'', U(p')] &\leq (L + 1)(\varepsilon + \operatorname{dist}[(x'', y''), K^{-1}(p')]) + |b'' - b'| \\ &\leq (L + 1)(\varepsilon + e[K^{-1}(p''), K^{-1}(p')]) + |b' - b''| \end{aligned}$$

whence

$$e[U(p''), U(p')] \leq (L + 1)\, e[K^{-1}(p''), K^{-1}(p')] + |b' - b''|. \tag{30}$$

A similar inequality holds by exchanging p' and p'', thus U is Lipschitz continuous on $\Omega(\delta)$ by (28), hence $M \times U$ is by Lemma 5.2, whence (7).

Proof of Theorem 4.3. By Theorem 4.1 (7) is equivalent to Lipschitz continuity of K^{-1} on $\Omega(\delta)$. By the boundedness assumption of $S(p)$ and (26), K^{-1} is locally bounded around 0. So we can apply [11, Theorem 5.11, (e)], taking into account that

$$w \in D^* K^{-1}(0, z)(0) \iff 0 \in D^* K(z)(-w)$$

(see [15, 8.(19)]).

Proof of Proposition 4.4. Consider $\bar{Q} = (\bar{f}, \bar{g})$ where

$$\bar{f}(x) = \frac{1}{4} \sum_{j=1}^{n} x_j^4; \quad \bar{g}_j(x) = -1, \quad j = 1, ..., q$$

and $\bar{\Omega}$ contains a sequence (a_j, b_j) with every $a_j \neq 0$ and $a_j \to 0$, moreover $|p| \leq \delta$ if $p \in \bar{\Omega}$. Then for all $p = (a, b)$ with $|p| \leq \delta$, $Q(p)$ consists of minimizing $\bar{f}(x) - a^T x$ on the whole R^n, hence the set of minimizers of $\bar{Q}(p)$ reduces to the singleton $\sqrt[3]{a} = (\sqrt[3]{a_1}, ..., \sqrt[3]{a_n})^T$ with multiplier 0, whence $\bar{Q} \in IC$. Following (6) we consider the Kojima functions K of Q, $\bar{K}$ of $\bar{Q}$ and

$$\begin{aligned} K(x, y) - \bar{K}(x, y) &= F(x, y) = \\ &= \left(\nabla f(x) + \sum_{j=1}^{q} y_j^+ \nabla g_j(x) - x^3, g_j(x) + 1, j = 1, ..., q \right) \end{aligned} \tag{31}$$

where x^3 denotes the vector of components $x_j^3, j = 1, ..., n$. Let $K(z) \in \Omega$, then $K(z) = p = (a, b)$ with $|p| \leq \delta$ and, by (A2) and Proposition 5.1,

$$z = (x, y) \ \text{ with } x \in M(p), \qquad y = u + g(x) - b, \tag{32}$$

hence $|x| \le s$ by (11). From the proof of [19, Proposition 1] we have

$$|u| \le \frac{|\nabla f(x) - a||x - x_0|}{\varepsilon - \delta}$$

where ε is given by (10) and x_0 by (A3). Thus

$$|u| \le k_1 = \left(\frac{s + |x_0|}{\varepsilon - \delta}\right) \max\{|\nabla f(x) - a| : |x| \le s, |a| \le \delta\}, \tag{33}$$

hence if $K(z) \in \Omega$ then

$$|z| \le k_2 = s + k_1 + \max\{|g(x) - b| : |x| \le s, |b| \le \delta\} \tag{34}$$

where k_1 is given by (33). Let now $\bar{K}(z) \in \Omega$, then again by Proposition 5.1 and (32), for some $(a, b) \in R^{n+q}$ we have $z = (\sqrt[3]{a}, -1 - b)$ where (abusing notation) 1 denotes the vector of R^q with all components equal to 1. It follows that $|z| \le n\sqrt[3]{\delta} + q + \delta$. Summarizing, D is contained in the closed ball of R^{n+q} with center 0 and radius

$$r = \max\{k_2, n\sqrt[3]{\delta} + q + \delta\}, \tag{35}$$

k_2 given by (34). Then

$$\begin{aligned} &\operatorname{dist}(Q, IC) \le \operatorname{dist}(Q, \bar{Q}) \le c_1 \\ &= \sup\left\{\frac{|F(z'') - F(z')|}{|z'' - z'|} : |z'| \le r, |z''| \le r\right\}, \end{aligned} \tag{36}$$

r given by (35) and F by (31).

Lemma 5.3. *Let Q fulfill* (A3) *and* (A4) *with Kojima function K. Suppose there exists a constant c^* such that*

$$\text{hausd}\,[K^{-1}(p''), K^{-1}(p')] \le c^*|p'' - p'|$$

for each sufficiently small p', p'' in the parameter set of Q. Then

$$\text{cond}\ Q \le (L + 2)c^* + 1. \tag{37}$$

Proof of Lemma 5.3. By Theorem 4.1, Q is well-conditioned. By (29)

$$\text{hausd}[M(p''), M(p')] \le \text{hausd}[K^{-1}(p''), K^{-1}(p')].$$

By (30)

$$\text{hausd}[U(p''), U(p')] \le (L + 1)\,\text{hausd}[K^{-1}(p''), K^{-1}(p')] + |p'' - p'|,$$

and

$$\text{hausd}[S(p''), S(p')] \le \text{hausd}[M(p''), M(p')] + \text{hausd}[U(p''), U(p')]$$

by (25). The conclusion follows by (4).

Proof of Proposition 4.5. By (37) of Lemma 5.3, we need to show that c_2 given by (16) is a Lipschitz constant of the set-valued mapping $K^{-1} : \Omega \to R^{n+q}$. Since K^{-1} is locally bounded on Ω by (17) and (26), we apply [11, Theorem 5.11, (5.9)]. The Lipschitz modulus of K^{-1} at $p = 0$ is then

$$(\operatorname{lip} K^{-1})(0) = \sup\{||D^* K^{-1}(0, z)|| : z \in K^{-1}(0)\}$$

where D^* denotes coderivative, see [15, Definition 8.3]. Of course K^{-1} is pseudo-Lipschitz around every $(0, z), z \in K^{-1}(0)$ (Theorem 4.1), hence by [11, Theorem 5.7, (5.8)] and [6, Theorem 4B.2, (2)] we get

$$(\operatorname{lip} K^{-1})(0) = \sup\{\limsup_{p \to 0, z' \to z, K(z') = p} |DK^{-1}(p, z')|^- : K(z) = 0\}$$

where D denotes graphical derivative, see [15, 8G]. By [6, p. 200] we have $DK^{-1}(p, z') = DK(z', p)^{-1}$, hence

$$(\operatorname{lip} K^{-1})(0) = \sup\{\limsup_{z' \to z} |DK(z')^{-1}|^- : K(z) = 0\}. \tag{38}$$

In (38), since K is single-valued, we write $DK(z')$ instead of $DK(z', p)$ with $K(z') = p$; by continuity of K, if $z' \to z$ then $p = K(z') \to K(z) = 0$. By [6, p. 202]

$$\begin{aligned} |DK(z')^{-1}|^- &= \sup_{|v| \le 1} \inf\{|w| : w \in DK(z')^{-1}(v)\} \\ &= \sup_{|v| \le 1} \inf\{|w| : v \in DK(z')(w)\}. \end{aligned} \tag{39}$$

Now we need

Lemma 5.4. *For any (x, y), the following are equivalent:*

$$v \in DK(x, y)(w); \tag{40}$$

$$v = A(x, y)(w_1, \alpha, \beta) \tag{41}$$

where $w = (w_1, w_2), w_1 \in R^n, w_2 \in R^q$, and α, β are given by (13), (14) *respectively.*

Taking Lemma 5.4 for granted, by (38), (39) and (15) we see that (lip K^{-1})(0) agrees with the right-hand side of (16), whence (18).

Proof of Lemma 5.4. We compute $DK(x, y)(w)$ making use of [15, 8(20)]. Write $z = (x, y)$ with $x \in R^n, y \in R^q$ and $w^* = (w_1^*, w_2^*)$ with $w_1^* \in R^n, w_2^* \in R^q$, and let $t > 0$. We consider the first n component of

$$\frac{1}{t}[K(z + tw^*) - K(z)] \tag{42}$$

which are given by

$$\frac{1}{t}\left[\nabla f(x+tw_1^*)+\sum_{i=1}^{q}(y_i+tw_{2i}^*)^+\nabla g_i(x+tw_1^*)-\nabla f(x)-\sum_{i=1}^{q}y_i^+\nabla g_i(x)\right] \tag{43}$$

as we obtain from (5). Moreover the remaining q components of (42) are given by

$$\frac{1}{t}\left[g_i(x+tw_1^*)-(y_i+tw_{2i}^*)^- - g_i(x)+y_i^-\right], i=1,...,q. \tag{44}$$

We compute the graphical derivative of the positive and negative part, and by letting $t \downarrow 0, w^* \to w = (w_1, w_2)$ we obtain from (43), after simple manipulations,

$$\left[\nabla^2 f(x)+\sum_{i=1}^{q}y_i^+\nabla^2 g_i(x)\right]w_1+\sum_{i=1}^{q}\alpha_i\nabla g_i(x), \tag{45}$$

and from (44)

$$\nabla g_i(x)^T w_1+\beta_i, \quad i=1,...,q \tag{46}$$

with α_i given by (13) and β_i by (14). Summarizing, $DK(x,y)(w)$ is single-valued for every x, y and w, and if $v = (v_1, v_2)$ with $v_1 \in R^n, v_2 \in R^q$, (40) is equivalent by (45) and (46) to

$$v_1=\left[\nabla^2 f(x)+\sum_{i=1}^{q}y_i^+\nabla^2 g_i(x)\right]w_1+\sum_{i=1}^{q}\alpha_i\nabla g_i(x),$$
$$v_{2i}=\nabla g_i(x)^T w_1+\beta_i, i=1,...,q$$

that is (41), as required.

Proof of Proposition 4.8. According to [11, Theorem 5.11, (5.9)] we show that T given by (20) agrees with

$$\sup\{||D^*K^{-1}(0,z)|| : z \in K^{-1}(0)\}$$

namely the Lipschitz modulus of K^{-1} at $p = 0$ (called in [11] the bound of local Lipschitzness for K^{-1} around 0). To this aim, by [15, 8(19)] we get

$$v \in D^*K^{-1}(0,z)(w) \iff -w \in D^*K(z)(-v)$$

for each z such that $K(z) = 0$. Hence by the definition of the norm (see [11, p. 23])

$$||D^*K^{-1}(0,z)|| = \sup\{|v| : -w \in D^*K(z)(-v), |w| \le 1\}.$$

By (5) and [15, 10.43 and 8.34], if $z = (x, y)$,

$$D^*K(x,y)(-v) = -E_1^T v + D^*E_2(-v)$$

where E_1, E_2 are as after (20), hence (20) is the Lipschitz modulus as claimed. By Theorem 4.1, this proves the first conclusion. The last follows by (37) of Lemma 5.3, since c_3 is a Lipschitz constant of K^{-1} near 0.

References

[1] D. Amelunxen, P. Bürgisser: A coordinate-free condition number for convex programming, SIAM J. Optim. 22 (2012) 1029–1041.

[2] M. Bianchi, G. Kassay, R. Pini: Conditioning for optimization problems under general perturbations, Nonlinear Anal. 75 (2013) 37–45.

[3] M. Bianchi, E. Miglierina, E. Molho, R. Pini: Some results on condition numbers in convex multiobjective optimization, Set-Valued Var. Anal. 21 (2013) 47–65.

[4] L. Blum, F. Cucker, M. Shub, S. Smale: Complexity and Real Computation, Springer, New York (1998).

[5] P. Bürgisser, F. Cucker: Condition, Springer, Berlin (2013).

[6] A. L. Dontchev, R. T. Rockafellar: Implicit Functions and Solution Mappings, Springer, New York (2009).

[7] J.-B. Hiriart-Urruty, C. Lemaréchal: Convex Analysis and Minimization Algorithms I, Springer, Berlin (1993).

[8] D. Klatte, B. Kummer: Nonsmooth Equations in Optimization, Kluwer, Dordrecht (2002).

[9] D. Klatte, B. Kummer: Aubin property and uniqueness of solutions in cone constrained optimization, Math. Meth. Oper. Res. 77 (2013) 291–304.

[10] O. L. Mangasarian, T.- H. Shiau: Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems, SIAM J. Control Optim. 25 (1987) 583–595.

[11] B. Mordukhovich: Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, Trans. Amer. Math. Soc. 340 (1993) 1–35.

[12] J. Renegar: Some perturbation theory for linear programming, Math. Program. 65 (1994) 73–91.

[13] J. Renegar: Linear programming, complexity theory and elementary functional analysis, Math. Program. 70 (1995) 279–351.

[14] S. M. Robinson: A characterization of stability in linear programming, Oper. Res. 25 (1977) 435–447.

[15] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).

[16] T. Zolezzi: On the distance theorem in quadratic optimization, J. Convex Analysis 9 (2002) 693–700.

[17] T. Zolezzi: Condition number theorems in optimization, SIAM J. Optim. 14 (2003) 507–516.

[18] T. Zolezzi: Condition number theorems in linear-quadratic optimization, Numer. Funct. Anal. Optim. 32 (2011) 1381–1404.

[19] T. Zolezzi: A condition number theorem in convex programming, Math. Program. 149 (2015) 195–207.