Mathematical Programming 49 (1990) 123-138
North-Holland

Implicit functions and sensitivity of stationary points

H.Th. Jongen*
Lehrstuhl C für Mathematik, RWTH Aachen, D-5100 Aachen, FR Germany

D. Klatte
Pädagogische Hochschule Halle-Kötchen, Sektion Mathematik und Physik, DDR-4050 Halle (Saale), GDR

K. Tammer
Technische Hochschule Leipzig, Sektion Mathematik und Informatik, DDR-7030 Leipzig, GDR

Received 1 February 1988
Revised manuscript received 10 January 1989

We consider the space $L(D)$ consisting of Lipschitz continuous mappings from D to the Euclidean n-space $\mathbb{R}^n$, D being an open bounded subset of $\mathbb{R}^n$. Let F belong to $L(D)$ and suppose that $\bar{x}$ solves the equation $F(x)=0$. In case that the generalized Jacobian of F at $\bar{x}$ is nonsingular (in the sense of Clarke, 1983), we show that for G near F (with respect to a natural norm) the system $G(x)=0$ has a unique solution, say $x(G)$, in a neighborhood of $\bar{x}$. Moreover, the mapping which sends G to $x(G)$ is shown to be Lipschitz continuous. The latter result is connected with the sensitivity of strongly stable stationary points in the sense of Kojima (1980); here, the linear independence constraint qualification is assumed to be satisfied.

Key words: Implicit function, stationary point, strong stability, Lipschitz continuity, generalized Jacobian, mapping degree.

1. Introduction

Given an open bounded subset D of the Euclidean n-space $\mathbb{R}^n$, we consider the space $L(D)$ consisting of all mappings $F: D \to \mathbb{R}^n$ which are Lipschitz continuous on D. One of the main results of this paper is an Implicit Function Theorem of the following type. Let $F \in L(D)$, $\bar{x} \in D$, suppose that $F(\bar{x})=0$ and, moreover, that the generalized Jacobian of F at $\bar{x}$ (in the sense of Clarke [1]) is nonsingular. Then there exist real positive numbers μ, ν, such that for each $G \in L(D)$ belonging to some μ-neighborhood of F (defined in a natural way), the system $G(x)=0$ has a unique solution, say $x(G)$, in the ν-neighborhood of $\bar{x}$; moreover, the mapping $G \mapsto x(G)$ is Lipschitz continuous.

* Honorary Professor, University of Hamburg, FR Germany.

As a basic application we shall show how the latter result can be used in order to analyse the sensitivity of stationary points in nonlinear programming problems of the type

$$P(f):\quad \text{minimize}\quad f_0(z)$$
$$\text{s.t.}\quad f_i(z)=0,\ i=1,\dots,l, \tag{1}$$
$$f_j(z)\le 0,\ j=l+1,\dots,m,$$

where $f_0, f_1, \dots, f_m$ are supposed to be twice continuously differentiable functions from $\mathbb{R}^n$ to $\mathbb{R}$.

In fact, following Kojima [7], the *Karush-Kuhn-Tucker* stationary conditions for $P(f)$ can be written as a system of equations

$$F(z, y)=0, \tag{2}$$

defined by

$$F(z,y)=\begin{pmatrix} \nabla f_0(z)+\sum_{i=1}^{l} y_i\nabla f_i(z)+\sum_{j=l+1}^{m} y_j^+\nabla f_j(z) \\ -f_1(z) \\ \vdots \\ -f_l(z) \\ y_{l+1}^- - f_{l+1}(z) \\ \vdots \\ y_m^- - f_m(z) \end{pmatrix} \tag{3}$$

where ∇f denotes the gradient of f and for $\alpha\in\mathbb{R}$ we put

$$\alpha^+=\max\{0,\alpha\},\qquad \alpha^-=\min\{0,\alpha\}. \tag{4}$$

Each solution $x=(z, y)$ of the equation $F(z, y)=0$ will be called a *stationary point* of $P(f)$.

In [7] Kojima introduced the concept of *strong stability* of stationary points, and presented necessary and sufficient conditions for it in terms of first and second order derivatives of the functions f_i, $i=0,1,\dots,m$.

Now, assuming the Linear Independence Constraint Qualification, we shall show that Kojima's conditions on strong stability are equivalent to the nonsingularity of the generalized Jacobian of F at the stationary point under consideration. The Lipschitz continuous dependence of stationary points on C^2-perturbations of the functions f_i, $i=0,1,\dots,m$, is then a consequence of our Implicit Function Theorem. In this way, our results give a deeper insight into the concept of strong stability and present a connection with Clarke's theory on nondifferentiable mappings (cf. [1]). We mention that a relationship with Robinson's concept of strong regularity, introduced in [13], is elaborated in [6].

The paper is organized as follows. In Section 2 we are concerned with the above mentioned Implicit Function Theorem. As an interesting by-product we state a result on the mapping degree of a Lipschitz continuous mapping. Then, Section 3 contains the relationship with strong stability of stationary points.

In the sequel we will use the symbols cl A, conv A and bd A which stand for the closure, the convex hull and the boundary of a set $A \subset \mathbb{R}^n$, respectively. Moreover, $\nabla^2 f$ denotes the Hessian of f. By $\|\cdot\|$ we denote the Euclidean norm on $\mathbb{R}^n$.

2. An Implicit Function Theorem

Using the notation of Section 1, we introduce a norm on the space $L(D)$ by

$$\|G\|_L := \sup_{x \in D} \max\{\|G(x)\|, \operatorname{Lip}(G)\}$$

for each $G \in L(D)$, where

$$\operatorname{Lip}(G) := \inf\{c \mid \|G(x) - G(y)\| \leq c\|x - y\| \text{ for all } x, y \in D\}.$$

Now we recall the definition and some properties of the generalized Jacobian of a Lipschitz continuous vector function (cf. [1]). Given $G \in L(D)$ and $\bar{x} \in D$, the set of (n, n)-matrices

$$\partial G(\bar{x}) := \operatorname{conv}\{M \mid \exists x^k \to \bar{x}\colon x^k \in E_G \text{ and } \nabla G(x^k) \to M\}$$

is called the *generalized Jacobian* of G at $\bar{x}$, where $E_G \subset D$ is the set of all points x for which the Jacobian $\nabla G(x)$ exists. This definition is justified by the fact that each $G \in L(D)$ is almost everywhere on D differentiable (Rademacher's Theorem, cf. [3]). As usual, the set $\partial G(\bar{x})$ will be regarded as a subset of the linear space $\mathbb{R}^{n \times n}$ which is endowed with the associated matrix norm

$$|||M||| := \max\{\|Mh\| \mid h \in \mathbb{R}^n, \|h\| = 1\}$$

for each $M \in \mathbb{R}^{n \times n}$. Then the following holds for $\bar{x} \in D$ (cf. [1]):

(a) $\partial G(x)$ is a nonempty convex compact subset of $\mathbb{R}^{n \times n}$;

(b) ∂G is upper semicontinuous at $\bar{x}$;

(c) if G_1, $G_2 \in L(D)$ then $G_1 + G_2 \in L(D)$ and $\partial(G_1 + G_2)(\bar{x})v \subset \partial G_1(\bar{x})v + \partial G_2(\bar{x})v$ holds for all $v \in \mathbb{R}^n$.

Remark 2.1. For $G \in L(D)$ we have the following inequality:

$$\sup_{\substack{M \in \partial G(x) \\ x \in D}} |||M||| \leq \operatorname{Lip}(G).$$

To show this, let $x \in D$ and $M \in \partial G(x)$ be arbitrarily chosen. By definition of $\partial G(x)$, there are matrices $M_1, \ldots, M_l$ and sequences $\{x^{ik}\}$, $i = 1, \ldots, l$, such that

$$M \in \operatorname{conv}\{M_1, \ldots, M_l\},$$

$$x^{ik} \in E_G \quad \forall k \ \ \forall i \in \{1, \ldots, l\} \qquad \text{and} \qquad x^{ik} \xrightarrow[k\to\infty]{} x \quad \forall i \in \{1, \ldots, l\},$$

$$M_i = \lim_{k\to\infty} \nabla G(x^{ik}) \quad \forall i \in \{1, \ldots, l\}.$$

Let ε be any positive real number. For each $i \in \{1, \ldots, l\}$ and each $k \in \{1, 2, \ldots\}$ we have

$$G(x^{ik} + \beta h) = G(x^{ik}) + \beta \nabla G(x^{ik})h + \mathrm{o}(\beta)$$

for all $h \in \mathbb{R}^n$ with $\|h\| = 1$ and all $\beta > 0$ (where $\mathrm{o}(\beta) \cdot \beta^{-1} \xrightarrow[\beta\to 0]{} 0$), since $x^{ik} \in E_G$ $\forall k$ $\forall i \in \{1, \ldots, l\}$. Hence

$$\begin{aligned} \|\nabla G(x^{ik})h\| &\leqslant \frac{1}{\beta} \|G(x^{ik} + \beta h) - G(x^{ik})\| + \frac{\mathrm{o}(\beta)}{\beta} \\ &\leqslant \operatorname{Lip}(G) + \varepsilon, \end{aligned}$$

if β is sufficiently small. Passing to the limit, we get

$$\|M_i h\| \leqslant \operatorname{Lip}(G) + \varepsilon \quad \forall i \in \{1, \ldots, l\},$$

and so

$$|||M_i||| \leqslant \operatorname{Lip}(G) \quad \forall i \in \{1, \ldots, l\} \qquad \text{and} \qquad |||M||| \leqslant \operatorname{Lip}(G),$$

where we have taken into account that ε was an arbitrarily chosen positive number.

Remark 2.2. For G_1, $G_2 \in L(D)$ satisfying $\|G_1 - G_2\|_L \leqslant \varepsilon$ (with $\varepsilon > 0$), one has

$$\partial G_1(x)v \subset \partial G_2(x)v + \varepsilon B_n \quad \forall x \in D, \ \forall v \in \mathbb{R}^n\colon \|v\| = 1,$$

where B_n is the closed unit ball in $\mathbb{R}^n$. To show this, we only note that for all $x \in D$ and all $v \in \mathbb{R}^n$ with $\|v\| = 1$,

$$\partial G_1(x)v \subset \partial G_2(x)v + \partial(G_1 - G_2)(x)v \subset \partial G_2(x)v + \operatorname{Lip}(G_1 - G_2)B_n,$$

where Property (c) of generalized Jacobians and Remark 2.1 were used.

Now we shall formulate the main result of this paper. We shall say that a nonempty subset S of $\mathbb{R}^{n\times n}$ is *nonsingular*, if every matrix $M \in S$ is nonsingular. In the following we shall denote by $\mathring{B}_n$ the open unit ball in $\mathbb{R}^n$, by $B(\bar{x}, \varepsilon)$ and $\mathring{B}(\bar{x}, \varepsilon)$ the closed and the open ε-neighborhood of $\bar{x}$, respectively, by $B_{n\times n}$ the closed unit ball in $\mathbb{R}^{n\times n}$, and by $U_\mu(F)$ (with $F \in L(D)$) the set $\{G \in L(D) \mid \|F - G\|_L < \mu\}$.

Theorem 2.1 (Implicit Function Theorem). *Let D be a nonempty open bounded subset of $\mathbb{R}^n$, let $F \in L(D)$, and let $\bar{x} \in D$ be a point satisfying $F(\bar{x}) = 0$. Suppose that $\partial F(\bar{x})$ is nonsingular. Then there exist positive real numbers ν, μ and γ such that the following holds:*

(i) *For each $G \in U_\mu(F)$, $B(\bar{x}, \frac{1}{2}\nu)$ contains a solution $x(G)$ of $G(x) = 0$ which is unique in $\mathring{B}(\bar{x}, \nu)$.*

(ii) *The mapping $G \mapsto x(G)$ satisfies a Lipschitz condition on $U_\mu(F)$ with Lipschitz modulus γ, i.e., if $G_1, G_2 \in U_\mu(F)$ then $\|x(G_1) - x(G_2)\| \leq \gamma \|G_1 - G_2\|_L$.*

It is worth noting that the proof of the theorem will in fact provide that

$$\|x(G_1) - x(G_2)\| \leq \gamma \max_{x \in V} \|G_1(x) - G_2(x)\| \quad \text{for all } G_1, G_2 \in U_\mu(F),$$

where $V := B(\bar{x}, \frac{1}{2}\nu)$.

The proof of Theorem 2.1 is based on the Inverse Function Theorem for Lipschitz functions, given by Clarke [1], and on a lemma on the degree of a Lipschitz continuous function Φ. First we shall present the above mentioned Inverse Function Theorem in a version which is convenient for our purposes.

Lemma 2.1 (Clarke [1]). *Let Q be a nonempty open subset of $\mathbb{R}^n$, let $\Phi : Q \to \mathbb{R}^n$ be Lipschitz continuous on Q, $x^0 \in Q$ and $\Phi(x^0) = 0$. Let Ω be a nonempty, convex and compact subset of $\mathbb{R}^{n \times n}$, and suppose that Ω is nonsingular. Further, suppose that there is a positive real number r such that $B(x^0, r) \subset Q$ and $\partial\Phi(x)v \subset \Omega v$ for all $x \in B(x^O, r)$ and all $v \in \mathbb{R}^n$ with $\|v\| = 1$. Let $\delta := \min\{\|Mv\| \mid M \in \Omega, \|v\| = 1\}$ (hence $\delta > 0$).*

Then there exists a mapping $x(\cdot) : \frac{1}{2}r\delta B_n \to B(x^0, r)$ satisfying the following properties:

(a) *For each $v \in \frac{1}{2}r\delta\mathring{B}_n$, $x(v)$ is a solution of the equation $\Phi(x) = v$, which is unique in $\mathring{B}(x^0, r)$.*

(b) *$x(\cdot)$ is Lipschitz continuous on $\frac{1}{2}r\delta\mathring{B}_n$ with Lipschitz modulus $1/\delta$.* □

Lemma 2.1 follows at once from Lemmas 1–3 in the proof of the Inverse Function Theorem in [1, Theorem 7.1.1]. We shall omit the details of the proof. For convenience, we have used in the formulation of the lemma a notation similar to that in [1, pp. 252–255]: replace the F in [1] by Φ and take the definition of Ω in [1, p. 254] into account.

Given a nonempty open bounded subset V of $\mathbb{R}^n$, a continuous function $\Phi : \operatorname{cl} V \to \mathbb{R}^n$ and a vector $c \in \mathbb{R}^n$ such that $\Phi(z) \neq c$ for each $z \in \operatorname{bd} V$, the Brouwer degree of Φ at c w.r.t. V, which we denote by $\deg(\Phi, V, c)$, is well defined (cf. [2, 10]).

Lemma 2.2. *Let Q be a nonempty open subset of $\mathbb{R}^n$. Let $\Phi : Q \to \mathbb{R}^n$ be Lipschitz continuous on Q, $x^0 \in Q$, and suppose that $\Phi(x^0) = 0$. If $\partial\Phi(x^0)$ is nonsingular, then there exists an $r > 0$ with $\mathring{B}(x^0, r) \subset Q$ such that*

$$\deg(\Phi, \mathring{B}(x^0, r), 0) = \operatorname{sign} \det A,$$

where A is any element from $\partial\Phi(x^0)$. In particular, we have $\deg(\Phi, \mathring{B}(x^0, r), 0) \in \{+1, -1\}$.

Proof. Let us first note that sign det A is constant and nonvanishing for all $A \in \partial\Phi(x^0)$, since $\partial\Phi(x^O)$ is nonsingular and connected. Now, choose $A \in \partial\Phi(x^0)$ and put

$$H(x, t) := (1-t)\Phi(x) + tA(x - x^0), \quad x \in \mathbb{R}^n, \ t \in \mathbb{R}.$$

We note

(i) $H(x^0, t) = 0$ for all $t \in \mathbb{R}$;

(ii) $\pi_x \partial H(x^0, t) \subset \partial\Phi(x^0)$ for all $t \in [0, 1]$,

where $\pi_x \partial H(x^0, t)$ signifies the set of all (n, n)-matrices M such that, for some vector $b \in \mathbb{R}^n$, the $(n, n+1)$-matrix $[M, b]$ belongs to $\partial H(x^0, t)$.

From (ii) we see that $\pi_x \partial H(x^0, t)$ is nonsingular for all $t \in [0, 1]$. Together with (i) and the Implicit Function Theorem of Clarke (cf. [1, p. 256]), we obtain: for each $\tilde{t} \in [0, 1]$ there exist $\tilde{\varepsilon} > 0$ and $\tilde{r} > 0$ such that $B(x^0, \tilde{r}) \subset Q$, and $(x^0, \tilde{t})$ is the unique solution of the equation $H(x, t) = 0$ in $B(x^0, \tilde{r}) \times (\tilde{t} - \tilde{\varepsilon}, \tilde{t} + \tilde{\varepsilon})$. Choose $\tilde{t}_i \in [0, 1]$, $i = 1, \ldots, k$ $(<\infty)$, such that $[0, 1] \subset \bigcup_{i=1}^{k} (\tilde{t}_i - \tilde{\varepsilon}_i, \tilde{t}_i + \tilde{\varepsilon}_i)$, with $(x^0, \tilde{t}_i)$ the unique solution of $H(x, t) = 0$ in $B(x^0, \tilde{r}_i) \times (\tilde{t}_i - \tilde{\varepsilon}_i, \tilde{t}_i + \tilde{\varepsilon}_i)$, and put $r := \min_{1 \le i \le k} \tilde{r}_i$. Then, $H(x, t) \neq 0$ for all $(x, t) \in \operatorname{bd} B(x^0, r) \times [0, 1]$. Hence, in virtue of the homotopy-invariance of the degree (cf. [2, Satz 1, p. 39]), we have

$$\deg(H(\cdot, t), \mathring{B}(x^0, r), 0) = \text{constant} \quad \text{for } t \in [0, 1].$$

Consequently, the assertions follow in view of the obvious fact that

$$\deg(x \mapsto A(x - x^0), \mathring{B}(x^0, r), 0) = \text{sign det } A. \qquad \square$$

As a corollary of Lemma 2.2, we immediately obtain:

Lemma 2.3. *Let Q be a nonempty open subset of $\mathbb{R}^n$. Let $\Phi : Q \to \mathbb{R}^n$ be Lipschitz continuous on Q, $x^0 \in Q$, and suppose that $\Phi(x^0) = 0$. Furthermore, suppose that $\partial\Phi(x^0)$ is nonsingular.*

Then, for each $\varepsilon > 0$ there is some $\alpha > 0$ such that for each continuous function $\tilde{\Phi} : Q \to \mathbb{R}^n$ with $\sup_{x \in Q} \|\Phi(x) - \tilde{\Phi}(x)\| < \alpha$, the ball $B(x^0, \varepsilon)$ contains at least one solution of $\tilde{\Phi}(x) = 0$. $\square$

Remark 2.3. We note that Lemma 2.3 may be considered as a special case of Kummer's Implicit Function Theorem [8, Theorem 4.1] for Kakutani mappings (which are introduced in [8] to be convex-valued and closed multifunctions $\Gamma : Z \subset \mathbb{R}^n \to \mathbb{R}^m$ such that for each $r > 0$ there is some $s > 0$ with $\Gamma(z) \cap sB_m \neq \emptyset$ if $z \in Z \cap rB_n$). The mentioned theorem particularly says that if a point $x^0 \in R^n$, and $\varepsilon > 0$ and a multifunction $\Gamma : B(x^0, \varepsilon) \to \mathbb{R}^n$ are given, and if for some $t > 0$ the multifunction $y \mapsto \Gamma^-(y) := \{x \in B(x^0, \varepsilon) \mid y \in \Gamma(x)\}$ defined on t. bd B_n is a Kakutani mapping, then the set $\{x \in B(x^0, \varepsilon) \mid 0 \in \tilde{\Gamma}(x)\}$ is nonempty whenever $\tilde{\Gamma}$ is a Kakutani mapping "near" Γ. With $\Gamma := \{\Phi\}$, $\tilde{\Gamma} := \{\tilde{\Phi}\}$, and by use of Lemma 2.1, it is not difficult to verify that Kummer's theorem applies to the situation in Lemma 2.3. We have preferred the approach via the degree lemma (Lemma 2.2), because this result is of interest in itself.

Now we prove Theorem 2.1. We recall the assumptions of Theorem 2.1: D is a nonempty open bounded subset of $\mathbb{R}^n$, $F: D \to \mathbb{R}^n$ is Lipschitz continuous on D, and $\bar{x} \in D$ is a point such that $F(\bar{x}) = 0$ holds, and, moreover, $\partial F(\bar{x})$ is nonsingular.

Proof of Theorem 2.1. First we shall introduce a triple (ν, η, δ) of positive real numbers which will be used to construct a triple (ν, μ, γ) such that (i) and (ii) in the statement of our theorem hold. Let ν, η, δ satisfy the following properties:

$$B(\bar{x}, 3\nu) \subset D, \tag{5}$$

$$\Omega := \partial F(\bar{x}) + 2\eta B_{n\times n} \text{ is nonsingular,} \tag{6}$$

$$\partial G(x)v \subset \Omega v \quad \forall G \in U_\eta(F),\ \forall x \in B(\bar{x}, 3\nu),\ \forall v\colon \|v\| = 1, \tag{7}$$

$$\delta := \min\{\|Mv\| \mid M \in \Omega,\ \|v\| = 1\} \quad (\text{hence } \delta > 0). \tag{8}$$

The existence of such a triple (ν, η, δ) is clear: (6) can be ensured because $\partial F(\bar{x})$ is nonsingular, (7) is a consequence of the upper semicontinuity of ∂F at $\bar{x}$ (which implies $\partial F(x) \subset \partial F(\bar{x}) + \eta B_{n\times n}$ for all $x \in B(\bar{x}, 3\nu)$ with some $\nu > 0$, and so, by Remark 2.2, as $G \in U_\eta(F)$, $\partial G(x)v \subset \partial F(\bar{x})v + 2\eta \mathring{B}_n \subset \Omega v$ for all $x \in B(\bar{x}, 3\nu)$ and all v: $\|v\| = 1$); (5) can then be guaranteed without loss of generality, and, finally, in (8) the number δ is well defined and positive because of (6). Further, let $\alpha(\nu)$ be a positive real number such that

$$B(\bar{x}, \tfrac{1}{2}\nu) \text{ contains at least one solution of } G(x) = 0 \text{ if } G \in U_{\alpha(\nu)}(F). \tag{9}$$

The existence of $\alpha(\nu)$ follows from Lemma 2.3.

Let G be any element of $L(D)$ satisfying $\|F - G\|_L < \min\{\eta, \alpha(\nu)\}$. Then, by (9), $B(\bar{x}, \frac{1}{2}\nu)$ contains at least one solution of $G(x) = 0$. Now let $x(G)$ be any solution of $G(x) = 0$ in $\mathring{B}(\bar{x}, \nu)$. With $Q := D$, $\Phi := G$, $x^0 := x(G)$ and with Ω as given in (6) and $r := 2\nu$, all assumptions of Lemma 2.1 are satisfied (note that $B(x(G), 2\nu) \subset B(\bar{x}, 3\nu)$ and take (5)-(8) into account). Hence, by Lemma 2.1, there exists a mapping $x_G(\cdot): \nu\delta\mathring{B}_n \to \mathring{B}(x(G), 2\nu)$ such that

$$x_G(v) \text{ is the unique solution of } G(x) = v \text{ in } \mathring{B}(x(G), 2\nu), \tag{10}$$

$$x_G(\cdot) \text{ is Lipschitz on } \nu\delta\mathring{B}_n \text{ with modulus } 1/\delta. \tag{11}$$

Since $x(G) \in \mathring{B}(\bar{x}, \nu)$ and $\mathring{B}(\bar{x}, \nu) \subset \mathring{B}(x(G), 2\nu)$, $x(G)$ is the unique solution of $G(x) = 0$ in $\mathring{B}(x, \nu)$. Thus, we have shown

$$B(\bar{x}, \tfrac{1}{2}\nu) \text{ contains a solution } x(G) \text{ of } G(x) = 0 \text{ which is unique in } \mathring{B}(\bar{x}, \nu) \tag{12}$$

and

$$x_G(0) = x(G). \tag{13}$$

Proof of (*i*). Define

$$\mu := \tfrac{1}{2}\min\{\eta, \alpha(\nu)\}.$$

Since (12) even holds for each $G \in U_{2\mu}(F)$, (i) is shown.

Proof of (*ii*). First let $G \in U_\mu(F)$ be fixed. Let $x(G)$ and $x_G(\cdot)$ satisfy (10)–(13). Further, let ε be a real number satisfying

$$0 < \varepsilon < \min\{\tfrac{1}{2}\nu, \nu\delta/\mathrm{Lip}(G)\}.$$

Setting $Q := D$, $\Phi := G$, $x^0 := x(G)$ and taking into account that $\partial G(x(G))$ is non-singular (because of (12), (7) and (6)), we see that Lemma 2.3 applies to this situation. Hence, we obtain: there exists an α, with $0 < \alpha \leq \mu$, such that

$$\tilde{G}(x) = 0 \text{ is solvable in } B(x(G), \varepsilon) \text{ for all } \tilde{G} \text{ with } \|G - G\|_L < \alpha.$$

Now we fix α with $0 < \alpha \leq \mu$, and choose any $\tilde{G} \in U_\alpha(G)$. Since, by (12), $x(G) \in B(\bar{x}, \frac{1}{2}\nu)$ and so for all $z \in B(x(G), \varepsilon)$,

$$\|z - \bar{x}\| \leq \|z - x(G)\| + \|x(G) - \bar{x}\| \leq \varepsilon + \tfrac{1}{2}\nu < \nu,$$

each solution z of $\tilde{G}(x) = 0$ in $B(x(G), \varepsilon)$ belongs to $\mathring{B}(\bar{x}, \nu)$. On the other hand, we have norm $\|F - \tilde{G}\|_L \leq \|F - G\|_L + \|G - \tilde{G}\|_L < 2\mu$, and so the arguments used above in the derivation of (12) also apply to $\tilde{G}$ (instead of G), and we have that

$$B(\bar{x}, \tfrac{1}{2}\nu) \text{ contains a solution } x(\tilde{G}) \text{ of } \tilde{G}(x) = 0 \text{ which is unique in } \mathring{B}(\bar{x}, \nu). \tag{12'}$$

Thus, the point $x(\tilde{G})$ given by (12′) is the unique solution of $\tilde{G}(x) = 0$ in $B(x(G), \varepsilon)$. By (5), (12) and (12′), one has $x(G)$, $x(\tilde{G}) \in D$, therefore

$$\|G(x(\tilde{G}))\| = \|G(x(\tilde{G})) - G(x(G))\| \leq \mathrm{Lip}(G)\|x(\tilde{G}) - x(G)\| \leq \mathrm{Lip}(G)\varepsilon < \nu\delta,$$

where $G(x(G)) = 0$ was used. Hence, $G(x(\tilde{G})) \in \nu\delta\mathring{B}_n$, and so, by (10),

$$x(\tilde{G}) \text{ is the unique solution of } G(x) = G(x(\tilde{G})) \text{ in } \mathring{B}(x(G), 2\nu).$$

The definition of $x_G(v)$ in (10) thus implies

$$x(\tilde{G}) = x_G(G(x(\tilde{G}))).$$

Taking $\tilde{G}(x(\tilde{G})) = 0$, (11) and (13) into account, we see that

$$\begin{aligned}\|x(G) - x(\tilde{G})\| &= \|x_G(0) - x_G(G(x(\tilde{G})))\| \\ &\leq \frac{1}{\delta}\|G(x(\tilde{G}))\| \\ &= \frac{1}{\delta}\|G(x(\tilde{G})) - \tilde{G}(x(\tilde{G}))\| \\ &\leq \frac{1}{\delta}\sup_{x \in D}\|G(x) - \tilde{G}(x)\| \\ &\leq \frac{1}{\delta}\|G - \tilde{G}\|_L\end{aligned}$$

This means, we have shown that for each $G \in U_\mu(F)$ there is some $\alpha := \alpha(G) > 0$ such that

$$\|x(G) - x(\tilde{G})\| \leq \frac{1}{\delta} \|G - \tilde{G}\|_L \quad \text{if } \tilde{G} \in U_\alpha(G). \tag{14}$$

Take now any pair G_1, $G_2 \in U_\mu(F)$ and consider

$$H(x, t) := (1-t)G_1(x) + tG_2(x) \quad \forall x \in D,\ \forall t \in [0, 1].$$

It is easy to show that for each $t \in [0, 1]$,

$$H(\cdot, t) \in U_\mu(F).$$

Hence, (i) holds, i.e., for each $t \in [0, 1]$ there exists some point $x(H(\cdot, t))$ which satisfies (12) and (14) (put there $G := H(\cdot, t)$). In particular, $x(G_1) := x(H(\cdot, 0))$ and $x(G_2) := x(H(\cdot, 1))$. Hence, for each $t \in [0, 1]$ there is some $\alpha = \alpha(t)$ such that (14) holds:

$$\|x(H(\cdot, t)) - x(H(\cdot, \tilde{t}))\| \leq \frac{1}{\delta} \|(t - \tilde{t})(G_2 - G_1)\|_L \quad \text{if } t - \alpha < \tilde{t} < t + \alpha, \tag{15}$$

where in (14) one has to put $G := H(\cdot, t)$ and $\tilde{G} := H(\cdot, \tilde{t})$. This defines a covering of $[0, 1]$ by open sets $\{B(t, \alpha(t)) \mid t \in [0, 1]\}$ and so there exist finitely many numbers $0 = t_1 < t_2 < \cdots < t_{N-1} < t_N = 1$ such that

$$\bigcup_{i=1}^{N} B(t_i, \alpha(t_i)) \supset [0, 1]$$

and, moreover,

$$I_i := (t_{i+1} - \alpha(t_{i+1}), t_i + \alpha(t_i)) \neq \emptyset \quad (\forall i \in \{1, \ldots, N-1\}).$$

Now we choose numbers $s_1, s_2, \ldots, s_{N-1}$ satisfying $s_i \in I_i$ $(i = 1, \ldots, N-1)$ and $t_1 < s_1 < t_2 < \cdots < s_{N-1} < t_N$ and apply (15) with $(t, \tilde{t}) := (t_i, s_i)$ and $(t, \tilde{t}) := (s_i, t_{i+1})$ for $i = 1, \ldots, N-1$:

$$\begin{aligned}
&\|x(G_1) - x(G_2)\| = \|x(H(\cdot, 0)) - x(H(\cdot, 1))\| \\
&\quad \leq \|x(H(\cdot, t_1)) - x(H(\cdot, s_1))\| + \|x(H(\cdot, s_1)) - x(H(\cdot, t_2))\| \\
&\qquad + \cdots + \|x(H(\cdot, t_{N-1})) - x(H(\cdot, s_{N-1}))\| + \|x(H(\cdot, s_{N-1})) - x(H(\cdot, t_N))\| \\
&\quad \leq \frac{1}{\delta} \left(\sum_{i=1}^{N-1} \|(s_i - t_i)(G_2 - G_1)\|_L + \sum_{i=1}^{N-1} \|(t_{i+1} - s_i)(G_2 - G_1)\|_L \right).
\end{aligned}$$

By definition of $\|G\|_L$ one has $\|tG\|_L = t\|G\|_L$ for $t \geq 0$. Thus, we then have

$$\begin{aligned}
\|x(G_1) - x(G_2)\| &\leq \frac{1}{\delta} \left(\sum_{i=1}^{N-1} (s_i - t_i + t_{i+1} - s_i) \right) \|G_2 - G_1\|_L \\
&= \frac{1}{\delta} t_N \|G_2 - G_1\|_L = \frac{1}{\delta} \|G_2 - G_1\|_L,
\end{aligned}$$

where $s_i > t_i > s_{i-1}$ $(i = 2, \dots, N-1)$, $s_1 > t_1 = 0$ and $t_N = 1$ were used. With $\gamma := 1/\delta$, (ii) is shown. □

3. Strong stability of stationary points

Let us return to the nonlinear programming problems of the type $P(f)$, as introduced in Section 1, where $f = (f_0, f_1, \dots, f_m)$. Let $\mathscr{C}^2$ denote the space of twice continuously differentiable mappings from $\mathbb{R}^n$ to $\mathbb{R}^{m+1}$. For a given $f \in \mathscr{C}^2$ the associated mapping $F: \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$, introduced in (3), is piecewise continuously differentiable (in the sense of Kojima [7]) and hence, locally Lipschitz continuous. A point $\bar{x} = (\bar{z}, \bar{y})$ at which F vanishes is called a *stationary point* of $P(f)$, whereas $\bar{z}$ is called a *stationary solution* and $\bar{y}$ an *associated multiplier.* Obviously, an associated multiplier $\bar{y}$ is unique if the Linear Independence Constraint Qualification (LICQ) is satisfied at $\bar{z}$.

LICQ: *The set* $\{\nabla f_i(\bar{z}) \mid f_i(\bar{z}) = 0,\ 1 \leq i \leq m\}$ *is linearly independent.*

In this section we shall study the relationship of *strong stability of stationary points* and the *nonsingularity of the generalized Jacobian* of F, under the assumption of LICQ.

For $f \in \mathscr{C}^2$ and a subset $V \subset \mathbb{R}^n$ we put

$$\mathscr{N}(f, V) = \sup_{0 \leq i \leq m} \sup_{z \in V} \max\{|f_i(z)|, \|\nabla f_i(z)\|, |||\nabla^2 f_i(z)|||\}.$$

Definition 3.1 (cf. [7]). Let $\bar{z}$ be a stationary solution of $P(f)$ for a given $f \in \mathscr{C}^2$. Then, $\bar{z}$ is *strongly stable* if for some $\delta^* > 0$ and each $\delta \in (0, \delta^*]$ there exists an $\alpha(\delta) > 0$ such that for each $g \in \mathscr{C}^2$ with $\mathscr{N}(f - g, B(\bar{z}, \delta^*)) < \alpha(\delta)$ the set $B(\bar{z}, \delta)$ contains a stationary solution of $P(g)$ which is unique in $B(\bar{z}, \delta^*)$. A stationary point $(\bar{z}, \bar{y})$ of $P(f)$ for which $\bar{z}$ is strongly stable is called strongly stable as well.

The concept of strong stability plays a central role in parametric optimization, homotopy methods, multilevel methods and statements on local convergence in nonlinear optimization (cf. [4, 5, 7, 9, 12]). Under LICQ, Kojima presented necessary and sufficient algebraic conditions for strong stability [7, Theorem 3.5 and Corollaries 3.6, 4.3]. We will actually use the latter equivalent conditions; but before stating them we need some additional notation.

At a stationary point $(\bar{z}, \bar{y})$ we define

$$I_0 = \{i \in \{l+1, \dots, m\} \mid f_i(\bar{z}) = 0\}, \tag{16}$$

$$I_+ = \{i \in I_0 \mid \bar{y}_i > 0\}, \tag{17}$$

$$C = \nabla^2 f_0 + \sum_{i=1}^{l} \bar{y}_i \nabla^2 f_i + \sum_{j=l+1}^{m} \bar{y}_j^+ \nabla^2 f_j|_{\bar{z}}.$$

Moreover, we put for $J \subset \{1, \dots, m\}$,

$$A(J) = (\nabla f_i(\bar{z}),\ i \in J),$$

and for I with $I_+ \subset I \subset I_0$ (*assuming for simplicity that* $I = \{l+1, \ldots, k\}$),

$$M(I) = \begin{pmatrix} C & A(I_e \cup I) & 0 \\ -A^T(I_e \cup I) & 0 & 0 \\ -A^T(\bar{I}) & 0 & E \end{pmatrix}, \tag{18}$$

where $I_e = \{1, \ldots, l\}$, $\bar{I} = \{l+1, \ldots, m\} \backslash I$ and E is the unit matrix of order $m-k$.

Lemma 3.1 [7]. *Let $(\bar{z}, \bar{y})$ be a stationary point of $P(f)$ for a given $f \in \mathscr{C}^2$. Suppose, moreover, that LICQ is satisfied at $\bar{z}$.*

Then, $(\bar{z}, \bar{y})$ is strongly stable if and only if

$$\text{sgn det } M(I) \textit{ is constant and nonvanishing}$$
$$\textit{for all } I \textit{ with } I_+ \subset I \subset I_0. \quad \square \tag{19}$$

We will prove the following two theorems.

Theorem 3.1. *Let $f \in \mathscr{C}^2$ and let F be the associated mapping introduced in* (3) *and let $(\bar{z}, \bar{y})$ be a stationary point of $P(f)$.*

The, the nonsingularity of ∂F at $(\bar{z}, \bar{y})$ is equivalent with (19).

Theorem 3.2. *Let $(\bar{x} = (\bar{z}, \bar{y})$ be a stationary point for $P(f)$ with $f \in \mathscr{C}^2$. Suppose that $\bar{x}$ is strongly stable and the LICQ is satisfied at $\bar{z}$. Let δ^* be chosen according to Definition* 3.1.

Then there exist positive real numbers $\bar{\nu}$, $\bar{\mu}$ and $\bar{\gamma}$ such that the following hold:

(i) *For each $g \in \mathscr{C}^2$ with $\mathcal{N}(f-g, B(\bar{z}, \delta^*)) < \bar{\mu}$ (the latter set being denoted by $\mathcal{O}_{\bar{\mu}} f$) the set $B(\bar{x}, \frac{1}{2}\bar{\nu})$ contains a stationary point of $P(g)$, say $\bar{x}(g) := (\bar{z}(g), \bar{y}(g))$, which is unique in $\mathring{B}(\bar{x}, \bar{\nu})$.*

(ii) *The mapping $g \mapsto \bar{x}(g)$ is Lipschitz continuous on $\mathcal{O}_{\bar{\mu}} f$ with Lipschitz modulus $\bar{\gamma}$, i.e., if g^1, $g^2 \in \mathcal{O}_{\bar{\mu}} f$, then*

$$\|\bar{x}(g^1) - \bar{x}(g^2)\| \leq \bar{\gamma} \mathcal{N}(g^1 - g^2, B(\bar{z}, \delta^*)).$$

For the proof of Theorem 3.1 the following technical lemma from linear algebra is crucial.

Lemma 3.2. *Let N_0 be a nonsingular (n, n)-matrix and $a^1, \ldots, a^k$, $b^1, \ldots, b^k \in \mathbb{R}^n$. Moreover, let $\mathcal{M}$ be the set of matrices of the form $N(I)$,*

$$N(I) = N_0 + \sum_{i+I} a^i b^{iT}, \quad \textit{with } I \subset \{1, \ldots, k\}.$$

Then, sgn det N = sgn det N_0 *for all $N \in$* conv $\mathcal{M}$ *if and only if* sgn det N = sgn det N_0 *for all $N \in \mathcal{M}$.*

Proof. Let us put $\mathcal{M} = \{N_0, N_1, \ldots, N_k\}$, with $K = 2^k - 1$ and $N_j = N_0 + \sum_{i \in I_j} a^i b^{i\mathrm{T}}$, I_j being a specific subset of $\{1, \ldots, k\}$. Since $\mathcal{M} \subset \operatorname{conv} \mathcal{M}$, one direction of the proof is trivial. *Hence, from now on, suppose that*

$$\operatorname{sgn} \det N = \operatorname{sgn} \det N_0 \quad \text{for all } N \in \mathcal{M}. \tag{$*$}$$

For an arbitrary $j \in \{0, 1, \ldots, K\}$, each index set $I \subset \{1, \ldots, k\}$ with $I \cap I_j = \emptyset$, and each $r \in \{1, \ldots, k\} \backslash (I \cup I_j)$ we show at first — by induction on $q := |I|$ — that the following two assertions hold (a pair (I, r) as above is called *admissible*):

(i) $\operatorname{sgn} \det (N_j + \sum_{i \in I} \mu_i a^i b^{i\mathrm{T}}) = \operatorname{sgn} \det N_0$,

(ii) $1 + \mu_r b^{r\mathrm{T}} \left(N_j + \sum_{i \in I} \mu_i a^i b^{i\mathrm{T}} \right)^{-1} a^r > 0$,

for all $\mu_i \in [0, 1]$ $(i \in I \cup \{r\})$.

The case $q = 0$. In this case we have $I = \emptyset$, and (i) follows from the assumption $(*)$. In order to show (ii) we use formula (2.20) in [11, p. 198] and obtain:

$$\det(N_j + \mu_r a^r b^{r\mathrm{T}}) = (1 + \mu_r b^{r\mathrm{T}} N_j^{-1} a^r) \cdot \det N_j.$$

If in particular, we put $\mu_r = 1$, we have $N_j + a^r b^{r\mathrm{T}} \in \mathcal{M}$. This implies, according to the assumption $(*)$, that $1 + b^{r\mathrm{T}} N_j^{-1} a^r > 0$, and, consequently, that $1 + \mu_r b^{r\mathrm{T}} N_j^{-1} a^r > 0$ for all $\mu_r \in [0, 1]$.

Now assume that (i) and (ii) are satisfied for each admissible pair (I_0, r_0) with $|I_0| \leq q_0$. we shall prove (i) and (ii) for an arbitrary admissible pair (I, r) with $|I| = q_0 + 1$.

Let i_0 belong to I and define $I_0 = I \backslash \{i_0\}$. Note that $|I_0| = q_0$. Further, let $\mu_i \in [0, 1]$ for $i \in I \cup \{r\}$. In order to show assertion (i) for (I, r) we again use formula (2.20) in [11, p. 198], and obtain (the induction assumption guarantees the invertibility which is used):

$$\det\left(N_j + \sum_{i \in I} \mu_i a^i b^{i\mathrm{T}} \right) = (1 + \mu_{i_0} b^{i_0\mathrm{T}} \tilde{N}^{-1} a^{i_0}) \cdot \det \tilde{N},$$

where

$$\tilde{N} = \left(N_j + \sum_{i \in I_0} \mu_i a^i b^{i\mathrm{T}} \right). \tag{20}$$

From the induction assumption (on the pair (I_0, I_0)) it then follows:

$$\operatorname{sgn} \det\left(N_j + \sum_{i \in I} \mu_i a^i b^{i\mathrm{T}} \right) = \operatorname{sgn} \det \tilde{N} = \operatorname{sgn} \det N_0,$$

and hence, assertion (i) is satisfied for I.

In order to prove assertion (ii) we apply formula (1.13) in [11, p. 190], and obtain for *fixed* $\mu_i \in [0,1]$ $(i \in I_0 \cup \{r\})$ and *variable* $\mu_{i_0} \in [0,1]$:

$$\lambda(\mu_{i_0}) := 1 + \mu_r b^{r\mathrm{T}}\left(N_j + \sum_{i\in I} \mu_i a^i b^{i\mathrm{T}}\right)^{-1} a^r$$
$$= 1 + \mu_r b^{r\mathrm{T}}\left[\tilde{N}^{-1} - \mu_{i_0} \cdot \frac{\tilde{N}^{-1} a^{i_0} b^{i_0\mathrm{T}} \tilde{N}^{-1}}{(1+\mu_{i_0} b^{i_0\mathrm{T}} \tilde{N}^{-1} a^{i_0})}\right] a^r,$$

where $\tilde{N}$ is defined in (20).

According to the induction assumption we have:

$$1 + \mu_p b^{p\mathrm{T}} \tilde{N}^{-1} a^p > 0, \quad p \in \{r, i_0\},$$
$$1 + \mu_r b^{r\mathrm{T}}\left(\bar{N}_j + \sum_{i\in I_0} \mu_i a^i b^{iT}\right)^{-1} a^r > 0,$$

where $\bar{N}_j := N_j + a^{i_0} b^{i_0\mathrm{T}}$ belongs to $\mathcal{M}$ since $i_0 \in \{1,\dots,k\}\setminus I_j$. This implies $\lambda(0) > 0$, $\lambda(1) > 0$ and, for $\mu_{i_0} \in [0,1]$:

$$\operatorname{sgn} \lambda(\mu_{i_0}) = \operatorname{sgn} \tilde{\lambda}(\mu_{i_0}),$$

where, $\tilde{N}$ as in (20),

$$\tilde{\lambda}(\mu_{i_0}) := \lambda(\mu_{i_0})[1 + \mu_{i_0} b^{i_0\mathrm{T}} \tilde{N}^{-1} a^{i_0}].$$

Note that $\tilde{\lambda}(\cdot)$ depends affinely on μ_{i_0}. So, we obtain $\operatorname{sgn} \lambda(\mu_{i_0}) = \operatorname{sgn} \tilde{\lambda}(\mu_{i_0}) = 1$ for $\mu_{i_0} \in [0,1]$ and hence, assertion (ii) for (I, r) follows.

Finally, the nontrivial assertion of the lemma follows from (i) by using the subsequent representation for an arbitrary element $N \in \operatorname{conv} \mathcal{M}$:

$$N = \sum_{j=1}^{K} \lambda_j N_j + \Big(1 - \sum_{j=1}^{K} \lambda_j\Big) N_0$$
$$= N_0 + \sum_{i=1}^{k} \left(\sum_{j:i\in I_j} \lambda_j\right) a^i b^{iT} = N_0 + \sum_{i=1}^{k} \mu_i a^i b^{iT}. \tag{21}$$

Since $\lambda_j \geq 0$ $(i = 1,\dots,K)$ and $\sum_{j=1}^{K} \lambda_i \leq 1$, the inequality $0 \leq \mu_i \leq 1$ $(i = 1,\dots,k)$ is satisfied in (21). □

Proof of Theorem 3.1. The mapping F in (3) is a continuous selection of the finite number of C^1-mappings F^I, $I \subset \{l+1,\dots,m\}$, where

$$F^I(z,y) = \begin{pmatrix} \nabla f_0(z) + \sum\limits_{i\in I\cup\{1,\dots,l\}} y_i \nabla f_i(z) \\ -f_i(z),\ i \in I \cup \{1,\dots,l\} \\ y_j - f_j(z),\ j \in \bar{I} \end{pmatrix},$$

with $\bar{I} = \{l+1,\dots,m\}\setminus I$.

In a neighborhood of the point $(\bar{z}, \bar{y})$ the mapping F is pieced together by means of the mappings F^I with $I_+ \subset I \subset I_0$ (none of them being redundant), and I_0, I_+ as in (16) and (17).

Note that $\nabla F^I(\bar{z}, \bar{y}) = M(I)$, cf. (18). Then, from the definition of a generalized Jacobian it follows that $\partial F(\bar{z}, \bar{y}) = \operatorname{conv} \mathcal{M}$, where

$$\mathcal{M} = \{M(I) \mid I_+ \subset I \subset I_0\}. \tag{22}$$

Hence, the nonsingularity of $\partial F(\bar{z}, \bar{y})$ implies (19). In order to prove the converse, we use Lemma 3.2 with $\mathcal{M}$ as in (22), $N_0 = M(I_+)$ and $k = |I_0 \backslash I_+|$. The required representation for elements of $\mathcal{M}$ follows from the relation

$$M(I_2) - M(I_1) = a^i b^{i\mathrm{T}},$$

where

$$I_1 = I_2 \backslash \{i\}, \qquad a^i = \begin{pmatrix} \nabla f_i(z) \\ -e^i \end{pmatrix},$$

e^i is the ith unit vector in $\mathbb{R}^m$, and b^i is the $(n+i)$th unit vector in $\mathbb{R}^{n+m}$. This completes the proof of the theorem. □

The following lemma will be used in the proof of Theorem 3.2. Recall the definition of $\|G\|_L$ and $\mathcal{N}(f, V)$ in the first lines of Section 2 and 3, respectively.

Lemma 3.3. *For $f, g \in \mathscr{C}^2$, let F, G denote the associated mappings accordings to* (3).

Then for any nonempty, open, bounded and convex subset $D \subset \mathbb{R}^{n+m}$ there is a number $\lambda > 0$ such that for every $f, g \in \mathscr{C}^2$,

$$\|F - G\|_L \leqslant \lambda \mathcal{N}(f - g, D_1), \tag{23}$$

where $D_1 = \pi_z D$ and π_z is the projection $(z, y) \mapsto z$ $(\in \mathbb{R}^n)$.

Proof. Let π_y be the projection $(z, y) \mapsto y$ $(\in \mathbb{R}^m)$, and put $D_2 = \pi_y D$. As further abbreviations we put

$$h = f - g, \qquad H = F - G,$$

$$\alpha_i = \sup_{z \in D_1} |h_i(z)|, \qquad \beta_i = \sup_{z \in D_1} \|\nabla h_i(z)\|, \qquad \gamma_i = \sup_{z \in D_1} \|\|\nabla^2 h_i(z)\|\|,$$

$i = 0, 1, \ldots, m$, and, finally,

$$\delta_0 = \max_{1 \leqslant i \leqslant m} \sup_{y \in D_2} |y_i|.$$

In view of the definition of $\|H\|_L$, we have to estimate $\sup_{(z,y) \in D} \|H(z, y)\|$ as well as Lip H in terms of $\mathcal{N}(h, D_1)$. Using the inequality $|\alpha^+| \leqslant |\alpha|$ we obtain, with some $z^0 \in \operatorname{cl} D_1$, some $y^0 \in \operatorname{cl} D_2$, and $\lambda = (2m+1) \max\{\delta_0, 2\}$:

$$\begin{aligned}
\sup_{(z,y) \in D} \|H(z, y)\| &= \max_{(z,y) \in \operatorname{cl} D} \|H(z, y)\| = \|H(z^0, y^0)\| \\
&\leqslant \left\| \nabla h_0(z^0) + \sum_{i=1}^{l} y_i^0 \nabla h_i(z^0) + \sum_{j=l+1}^{m} (y_j^0)^+ \nabla h_j(z^0) \right\| + \sum_{i=1}^{m} |h_i(z^0)| \\
&\leqslant \|\nabla h_0(z^0)\| + \sum_{i=1}^{m} |y_i^0| \cdot \|\nabla h_i(z^0)\| + \sum_{i=1}^{m} |h_i(z^0)| \\
&\leqslant \beta_0 + \delta_0 \sum_{i=1}^{m} \beta_i + \sum_{i=1}^{m} \alpha_i \leqslant \lambda \mathcal{N}(h, D_1).
\end{aligned}$$

Note that β_i and γ_i are Lipshitz moduli for h_i and ∇h_i, respectively (with respect to D_1). Then using the inequalities $|\alpha^+| \leq |\alpha|$, $|\alpha^+ - \beta^+| \leq |\alpha - \beta|$, the following holds for every (z^1, y^1), $(z^2, y^2) \in D$:

$$\begin{aligned}
&\|H(z^1, y^1) - H(z^2, y^2)\| \\
&\leq \|\nabla h_0(z^1) - \nabla h_0(z^2)\| + \sum_{i=1}^m |y_i^1 - y_i^2| \cdot \|\nabla h_i(z^1)\| \\
&\quad + \sum_{i=1}^m |y_i^2| \cdot \|\nabla h_i(z^1) - \nabla h_i(z^2)\| + \sum_{i=1}^m |h_i(z^1) - h_i(z^2)| \\
&\leq \gamma_0 \|z^1 - z^2\| + \sum_{i=1}^m \beta_i |y_i^1 - y_i^2| + \delta_0 \sum_{i=1}^m \gamma_i \|z^1 - z^2\| + \sum_{i=1}^m \beta_i \|z^1 - z^2\| \\
&\leq \left(\gamma_0 + 2 \sum_{i=1}^m \beta_i + \delta_0 \sum_{i=1}^m \gamma_i\right) \cdot \left\| \begin{pmatrix} z_0^1 \\ y^1 \end{pmatrix} - \begin{pmatrix} z^2 \\ y^2 \end{pmatrix} \right\| \\
&\leq \lambda \cdot \mathcal{N}(h, D_1) \cdot \left\| \begin{pmatrix} z^1 \\ y^1 \end{pmatrix} - \begin{pmatrix} z^2 \\ y^2 \end{pmatrix} \right\|.
\end{aligned}$$

This implies the inequality $\operatorname{Lip}(H) \leq \lambda \mathcal{N}(h, D_1)$, and, consequently, (23) is proved. □

Proof of Theorem 3.2. Lemma 3.1 and Theorem 3.1 ensure that ∂F is nonsingular at $\bar{x} = (\bar{z}, \bar{y})$, where F is the mapping introduced in (3). Then, with $D := \mathring{B}(\bar{x}, \delta^*)$, Theorem 2.1 and Lemma 3.3 provide the desired results. □

Remark 3.1. A counterexample in [14, p. 219] shows that $z(\cdot)$ in Theorem 3.2 need *not* be *Lipschitz* continuous if LICQ is replaced by the weaker Mangasarian-Fromovitz Constraint Qualification.

Acknowledgement

We would like to thank the two anonymous referees for many valuable suggestions and remarks.

References

[1] F.M. Clarke, *Optimization and Nonsmooth Analysis* (Wiley, New York, 1983).
[2] K. Deimling, *Nichtlineare Gleichungen und Abbildungsgrade* (Springer, Berlin-Heidelberg-New York, 1974).
[3] H. Federer, *Geometric Measure Theory* (Springer, Berlin-Heidelberg-New York, 1969).
[4] J. Guddat, Hj. Wacker and W. Zulehner, "On imbedding and parametric optimization", *Mathematical Programming Study* 21 (1984) 79-96.

[5] H.Th. Jongen, T. Moebert and K. Tammer, "On iterated minimization in nonconvex optimization," *Mathematics of Operations Research* 11 (1986) 679-691.

[6] H.Th. Jongen, T. Moebert, J. Rueckmann and K. Tammer, "On intertia and Schur complement in optimization," *Linear Algebra and its Applications* 95 (1987) 97-109.

[7] M. Kojima, "Strongly stable stationary solutions in nonlinear programs," in: S.M. Robinson, ed., *Analysis and Computation of Fixed Points* (Academic Press, New York, 1980) pp. 93-138.

[8] B. Kummer, "Linearly and nonlinearly perturbed optimization problems," in: J. Guddat, H.Th. Jongen, B. Kummer and F. Nožička, eds., *Parametric Optimization and Related Topics* (Akademie Verlag, Berlin, 1987) pp. 249-267.

[9] R. Lehmann, "On the numerical feasibility of continuation methods for nonlinear programming problems," *Mathematische Operationsforschung und Statistik, Series Optimization* 15 (1984) 517-520.

[10] J.M. Ortega and W. Rheinboldt, *Iterative Solutions of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[11] D.V. Ouelette, "Schur complements and statistics," *Linear Algebra and its Applications* 36 (1981) 187-295.

[12] C. Richter, "Ein implementierbares Einbettungsverfahren der nichtlinearen Optimierung," *Mathematische Operationsforschung und Statistik, Series Optimization* 15 (1984) 545-553.

[13] S.M. Robinson, "Strongly regular generalized equations," *Mathematics of Operations Research* 5 (1980) 43-62.

[14] S.M. Robinson, "Generalized equations and their solutions, Part II: Applications to nonlinear programming," *Mathematical Programming Study* 19 (1982) 200-221.