DIRECTIONAL REGULARITY AND METRIC REGULARITY*

A. V. Arutyunov[1], E. R. Avakov[2] and A. F. Izmailov[3]

September 20, 2006

ABSTRACT

For general constraint systems in Banach spaces, we present the directional stability theorem based on the appropriate generalization of directional regularity condition, suggested earlier in [1]. This theorem contains Robinson's stability theorem but does not reduce to it. Furthermore, we develop the related concept of directional metric regularity which is stable subject to small Lipschitzian perturbations of the constraint mapping, and which is equivalent to directional regularity for sufficiently smooth mappings. Finally, we discuss some applications in sensitivity theory.

Key words. Metric regularity, Robinson's constraint qualification, directional regularity, directional metric regularity, feasible arc, sensitivity.
AMS subject classifications. 49K27, 49K40, 90C31.

---

* Research of the first two authors are supported by the Russian Foundation for Basic Research Grants 05-01-00193 and 05-01-00275, and by RF President's Grants NS-5344.2006.1 and NS-5813.2006.1 for the state support of leading scientific schools. The third author is supported by the Russian Foundation for Basic Research Grant 04-01-00341, by RF President's Grant NS-9394.2006.1 for the state support of leading scientific schools, and by RF President's Grant MD-2723.2005.1 for the state support of young doctors of sciences.
[1] Peoples Friendship University, Miklukho-Maklaya Str. 6, 117198 Moscow, Russia.
Email: arutun@orc.ru
[2] Institute for Control Problems RAS, Profsoyuznaya Str. 65, 117806 Moscow, Russia.
Email: era@maxmin.ru
[3] Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Operations Research, Leninskiye Gori, GSP-2, 119992 Moscow, Russia.
Email: izmaf@ccas.ru

1 Introduction. Directional Regularity

Let Σ be a topological space, X and Y be Banach spaces, and Q be a fixed closed set in Y. Consider a smooth mapping $F : \Sigma \times X \to Y$ (our smoothness hypotheses will be specified below), and set

$$D(\sigma) = \{x \in X \mid F(\sigma, x) \in Q\} \tag{1.1}$$

with $\sigma \in \Sigma$ playing the role of a parameter. For a given (base) parameter value $\sigma_0 \in \Sigma$, fix $x_0 \in D(\sigma_0)$. In this paper we are concerned with the following question: for which $(\sigma, x) \in \Sigma \times X$ close to (σ_0, x_0), and under which assumptions $\operatorname{dist}(x, D(\sigma))$ can be estimated from above via the "residual" of constraints in (1.1), that is, via $\operatorname{dist}(F(\sigma, x), Q)$? Here $\operatorname{dist}(z, S) = \inf_{s \in S} \|z - s\|$ stands for the distance from a point z to a set S.

The answer to this question is well-known provided Q is convex and the so-called Robinson's constraint qualification (CQ) is satisfied at x_0 for the mapping $F(\sigma_0, \cdot)$, that is,

$$0 \in \operatorname{int}\left(F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - Q\right), \tag{1.2}$$

where $\operatorname{int} S$ is the interior of a set S, and $\operatorname{im} \Lambda$ is the range (image space) of a linear operator Λ. According to Robinson's stability theorem [21] (see also [3, Theorem 2.87]), under these assumptions there exists a constant $c > 0$ such that the estimate

$$\operatorname{dist}(x, D(\sigma)) \leq c \operatorname{dist}(F(\sigma, x), Q) \tag{1.3}$$

holds *for all* $(\sigma, x) \in \Sigma \times X$ close enough to (σ_0, x_0).

In its turn, estimate (1.3) serves as a motivation for the very important concept of metric regularity. Apparently, the term "metric regularity" appeared for the first time in [4] but the concept dates back to earlier works [14, 20, 10] (or even to classical works [15, 12]; see also [6, 5, 19]), and it finds multiple applications in modern variational analysis. Specifically, the mapping $F : X \to Y$ is said to be metrically regular at $x_0 \in F^{-1}(Q)$ with respect to Q if there exists a constant $c > 0$ such that the estimate

$$\operatorname{dist}(x, F^{-1}(Q + y)) \leq c \operatorname{dist}(F(x) - y, Q) \tag{1.4}$$

holds for all $(x, y) \in X \times Y$ close enough to $(x_0, 0)$. Note that (1.4) is nothing else but the estimate (1.3) for $F(\sigma, x) = F(x) - \sigma$ and $\sigma = y$, i.e., for the special parametrization of the mapping F in question (the "right-hand side" perturbations). Thus, by Robinson's stability theorem, if Q is convex then Robinson's CQ

$$0 \in \operatorname{int}(F(x_0) + \operatorname{im} F'(x_0) - Q)$$

implies metric regularity of F at x_0 with respect to Q. Moreover, as is well-known (see, e.g., [3, Proposition 2.89]), under the appropriate smoothness hypothesis, the converse implication is true as well, and thus, metric regularity and Robinson's CQ are actually equivalent.

For more recent developments and extensions of the metric regularity theory see, e.g., [17, 22, 13, 18, 16] and references therein.

In particular, if Robinson's CQ does not hold, one cannot expect a smooth mapping to be metrically regular. Accordingly, for a parametric mapping F, estimate (1.3) for all (σ, x) close enough to (σ_0, x_0) cannot be guaranteed if (1.2) does not hold. However, we demonstrate below that under the regularity condition weaker than (1.2), estimate (1.3) is still valid but possibly not *for all* (σ, x) in a neighborhood of (σ_0, x_0): the set of appropriate (σ, x) will be specified. To this end, we give the following

Definition 1.1 The mapping $F(\sigma_0, \cdot) : X \to Y$ is *regular at* $x_0 \in D(\sigma_0)$ *in a direction* $\bar{y} \in Y$ if

$$0 \in \operatorname{int}\left(F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - \operatorname{cone}\{\bar{y}\} - Q\right), \tag{1.5}$$

where $\operatorname{cone} S$ stands for the conic hull of a set S.

Note that for $\bar{y} = 0$, condition (1.5) reduces to Robinson's CQ (1.2). Moreover, if the latter is satisfied, the directional regularity condition (1.5) holds in any direction $\bar{y} \in Y$, including $\bar{y} = 0$.

Condition (1.5) and the corresponding directional stability result were first suggested in [1] for the case of finite-dimensional Y. However, the estimate obtained in [1, Theorem 4.1] is somewhat weaker than (1.3). This is a consequence of the general framework adopted in [1]. Specifically, the authors first consider the case of equality constraints and direct set constraints with a closed convex set P, and prove the directional stability theorem with the estimate to the solution set only from points in P. Then they reduce (1.1) to this setting. On the other hand, the proof of directional stability theorem in [1] is very concise and clear, and in particular, it does not appeal to any set-valued analysis. At the same time, the assumption $\dim Y < \infty$ cannot be dropped in that proof (and hence, in all the results obtained in [1]) because the argument there employs (completely finite-dimensional) Brouwer's fixed point theorem. (We note, however, that in [1, Theorem 4.1], X can actually be just a normed linear space, not necessarily complete.)

In Section 2, we prove the directional stability theorem (Theorem 2.1) under the same set of assumptions as in [1], but with the resulting estimate of the "proper" form (1.3), and for a (possibly infinite-dimensional) Banach space Y. In particular, Theorem 2.1 contains Robinson's stability theorem but does not reduce to it, in general.

Furthermore, in Section 3, for a nonparametric mapping, we develop the directional metric regularity concept suggested by Theorem 2.1. In Theorem 3.1 we demonstrate that this property is stable subject to small Lipschitzian perturbations of F. This result combined with Theorem 2.1 implies the equivalence of directional regularity and directional metric regularity for sufficiently smooth mappings.

Finally, in Section 4, we demonstrate that Theorem 2.1 can be used in order to directly obtain various stability results, widely used in sensitivity analysis [3]. Specifically, assuming that Σ is a normed linear space, we consider the case when for a given direction $d \in \Sigma$ it holds that

$$0 \in \operatorname{int}\left(F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) + \operatorname{cone}\left\{\frac{\partial F}{\partial \sigma}(\sigma_0, x_0)d\right\} - Q\right). \tag{1.6}$$

Note that (1.6) is a particular case of (1.5) for a specific $\bar{y}$, namely for $\bar{y} = -\frac{\partial F}{\partial \sigma}(\sigma_0, x_0)d$. On the other hand, (1.5) can be interpreted as (1.6) with F replaced by $F(\sigma, y, x) = F(\sigma, x) - y$, where $y \in Y$ is regarded as an additional parameter, and with $d = (0, \bar{y}) \in \Sigma \times Y$.

In the context of mathematical programming problems, (1.6) is known as Gollan's condition [11]. It was extended to the general case in [2] (see also [3, Theorem 4.9]). Moreover, in parametric optimization, *this* condition (which is a particular case of (1.5)) is commonly known as the directional regularity condition. Taking into account the relations between the two conditions discussed above, the authors prefer to use the same name for the property stated in Definition 1.1. Note, however, that unlike (1.6), (1.5) does not depend on a specific parametrization at all: it is entirely a property of the unperturbed constraints. This makes our directional regularity particulary useful for unification of some diverse developments, like those based on Robinson's CQ and on customary directional regularity (1.6).

2 Directional Stability Theorem

In the sequel, we shall need some equivalent formulations of the directional regularity condition introduced in Definition 1.1.

Proposition 2.1 *Let Q be closed and convex.*

Then condition (1.5) *is equivalent to either of the following three conditions:*

$$\operatorname{cone}\{\bar{y}\} \cap \operatorname{int}\left(F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - Q\right) \neq \emptyset, \tag{2.1}$$

$$\bar{y} \in \operatorname{int}\left(\operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - R_Q(F(\sigma_0, x_0))\right), \tag{2.2}$$

and

$$\operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - \operatorname{cone}\{\bar{y}\} - R_Q(F(\sigma_0, x_0)) = Y, \tag{2.3}$$

where $R_S(z) = \operatorname{cone}(S - z)$ stands for the radial cone to a set S at a point $z \in S$.

Note that condition (2.1) can be expressed in the following form: there exists $\theta \geq 0$ such that

$$\theta\bar{y} \in \operatorname{int}\left(F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - Q\right). \tag{2.4}$$

Proof. (1.5) $\Rightarrow$ (2.1). The proof of this implication is almost identical to that of the corresponding assertion in [3, Theorem 4.9] (see the argument showing that (4.12) implies (4.13)). Define the multifunction $\Psi : X \times \mathbf{R} \to 2^Y$,

$$\Psi(x, \theta) = \begin{cases} F(\sigma_0, x_0) + \frac{\partial F}{\partial x}(\sigma_0, x_0)x - \theta\bar{y} - Q & \text{if } \theta \geq 0, \\ \emptyset & \text{if } \theta = 0. \end{cases} \tag{2.5}$$

Evidently, Ψ is a closed convex multifunction (that is, graph Ψ is a closed convex set; see, e.g., [3, p. 55]), and

$$\Psi(X \times \mathbf{R}) = F(\sigma_0, x_0) + \operatorname{im}\frac{\partial F}{\partial x}(\sigma_0, x_0) - \operatorname{cone}\{\bar{y}\} - Q,$$

and thus, (1.5) means that $0 \in \operatorname{int} \Psi(X \times \mathbf{R})$. Furthermore, $0 \in \Psi(0, 0)$, and hence, by the generalized open mapping theorem [3, Theorem 2.70] and by (2.5), $0 \in \operatorname{int} \Psi(X \times [0, 1])$. This means that there exists $\delta > 0$ such that

$$\begin{aligned} B_\delta(0) &\subset \Psi(X \times [0, 1]) \\ &= F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - \{\theta \bar{y} \mid \theta \in [0, 1]\} - Q, \end{aligned} \tag{2.6}$$

where $B_\delta(z)$ stands for the ball centered at z and of radius δ.

Fix $\tilde{\delta} > 0$ small enough so that $\tilde{\delta}\bar{y} \in B_\delta(0)$. Then inclusion (2.6) implies that there exists $\tilde{\theta} \in [0, 1]$ such that

$$\tilde{\delta}\bar{y} \in F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - \tilde{\theta}\bar{y} - Q,$$

and hence,

$$(\tilde{\delta} + \tilde{\theta})\bar{y} \in F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - Q.$$

The set in the right-hand side of the latter inclusion is convex and contains 0, and thus

$$\{(\tilde{\delta} + \tilde{\theta})\theta\bar{y} \mid \theta \in [0, 1]\} \subset F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - Q.$$

Then inclusion (2.6) implies that

$$\begin{aligned} B_\delta(0) &\subset F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - \bar{y} + \{\theta \bar{y} \mid \theta \in [0, 1]\} - Q \\ &\subset (1 + 1/(\tilde{\delta} + \tilde{\theta})) \left(F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - Q \right) - \bar{y}. \end{aligned}$$

It follows that

$$\begin{aligned} B_{\delta\theta}(\theta\bar{y}) &= \theta\bar{y} + B_{\delta\theta}(0) \\ &\subset F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - Q \end{aligned}$$

holds with $\theta = (1 + 1/(\tilde{\delta} + \tilde{\theta}))^{-1} > 0$, and (2.4) (and hence (2.1)) is thus proved.

(2.1) $\Rightarrow$ (2.2). Since $Q - F(\sigma_0, x_0) \subset R_Q(F(\sigma_0, x_0))$, condition (2.1) clearly implies that

$$\operatorname{int} \left(\operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - R_Q(F(\sigma_0, x_0)) \right) \neq \emptyset.$$

Suppose that (2.2) does not hold. Then by the first separation theorem [3, Theorem 2.13], there exists $\mu \in Y^*$ such that

$$\langle \mu, \bar{y} \rangle \leq \langle \mu, \eta \rangle \quad \forall\, \eta \in \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - R_Q(F(\sigma_0, x_0)).$$

This evidently implies that

$$\langle \mu, \theta\bar{y} \rangle \leq 0 \leq \langle \mu, y \rangle \quad \forall\, \theta \geq 0,\ \forall\, y \in F(\sigma_0, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0, x_0) - Q,$$

where the inclusion $Q - F(\sigma_0, x_0) \subset R_Q(F(\sigma_0, x_0))$ was again taken into account. Hence, μ separates $\text{cone}\{\bar{y}\}$ and $F(\sigma_0, x_0) + \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - Q$, and according to the first separation theorem [3, Theorem 2.13], this contradicts (2.1).

(2.2) $\Rightarrow$ (2.3). By (2.2), there exists $\delta > 0$ such that

$$\begin{aligned} \bar{y} + B_\delta(0) &= B_\delta(\bar{y}) \\ &\subset \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - R_Q(F(\sigma_0, x_0)), \end{aligned}$$

and hence

$$\begin{aligned} B_\delta(0) &\subset \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \bar{y} - R_Q(F(\sigma_0, x_0)) \\ &\subset \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} - R_Q(F(\sigma_0, x_0)). \end{aligned}$$

Thus,

$$0 \in \text{int}\left(\text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} - R_Q(F(\sigma_0, x_0))\right),$$

holds, which evidently implies (2.3).

(2.3) $\Rightarrow$ (1.5). The proof of this implication is almost identical to that of the corresponding assertion in [3, Proposition 2.95] (see the argument showing that (2.180) implies (2.178)). Define the multifunction $\Psi : X \times \mathbf{R} \times \mathbf{R} \to 2^Y$,

$$\Psi(x, \theta, \tau) = \begin{cases} \frac{\partial F}{\partial x}(\sigma_0, x_0)x - \theta\bar{y} - \tau(Q - F(\sigma_0, x_0)) & \text{if } \theta \geq 0,\ \tau \geq 0 \\ \emptyset & \text{otherwise.} \end{cases} \tag{2.7}$$

Evidently, Ψ is a closed convex multifunction, and

$$\Psi(X \times \mathbf{R} \times \mathbf{R}) = \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} - R_Q(F(\sigma_0, x_0)),$$

and thus, (2.3) implies that $0 \in \text{int}\,\Psi(X \times \mathbf{R} \times \mathbf{R})$. Furthermore, $0 \in \Psi(0, 0, 0)$, and hence, by the generalized open mapping theorem [3, Theorem 2.70] and by (2.7),

$$0 \in \text{int}\,\Psi(X \times \mathbf{R}_+ \times [0, 1]). \tag{2.8}$$

On the other hand,

$$\begin{aligned} \Psi(X \times \mathbf{R}_+ \times [0, 1]) &= \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} - \{\tau(q - F(\sigma_0, x_0)) \mid \tau \in [0, 1],\ q \in Q\} \\ &= F(\sigma_0, x_0) + \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} \\ &\quad -\{\tau q + (1 - \tau)F(\sigma_0, x_0)) \mid \tau \in [0, 1],\ q \in Q\} \\ &\subset F(\sigma_0, x_0) + \text{im}\,\frac{\partial F}{\partial x}(\sigma_0, x_0) - \text{cone}\{\bar{y}\} - Q, \end{aligned}$$

where the convexity of Q was taken into account. It follows that (2.8) implies (1.5). ∎

We shall also need the following

Proposition 2.2 *Let Q be convex, $y_0 \in Q$.*

Then for any $\bar{y} \in Y$ and any $\delta_1 > 0$, $\delta_2 > 0$ there exists $\varepsilon > 0$ and $\delta > 0$ such that

$$(Q - \operatorname{cone} B_\delta(\bar{y})) \cap B_\varepsilon(y_0) \subset Q \cap B_{\delta_1}(y_0) - \operatorname{cone} B_{\delta_2}(\bar{y}). \tag{2.9}$$

Proof. First suppose that $\bar{y} \in T_Q(y_0)$, where $T_Q(y_0) = \operatorname{cl} R_Q(y_0)$ is the tangent cone to Q at y_0. We claim that in this case

$$y_0 \in \operatorname{int}(Q \cap B_{\delta_1}(y_0) - \operatorname{cone} B_{\delta_2}(\bar{y})), \tag{2.10}$$

and hence, (2.9) evidently holds with an arbitrary $\delta > 0$ and a sufficiently small $\varepsilon > 0$.

Indeed, the interior of the set in the right-hand side of (2.10) is nonempty, and if (2.10) does not hold then by the first separation theorem [3, Theorem 2.13] there exists $\mu \in Y^*$ such that

$$\langle \mu, y \rangle \geq \langle \mu, y_0 \rangle \quad \forall\, y \in Q \cap B_{\delta_1}(y_0) - \operatorname{cone} B_{\delta_2}(\bar{y}). \tag{2.11}$$

Then evidently

$$\langle \mu, \eta \rangle \geq 0 \quad \forall\, \eta \in T_Q(y_0). \tag{2.12}$$

On the other hand, for any $y \in B_{\delta_2}(0)$ such that $\langle \mu, y \rangle > 0$ from (2.11) we obtain

$$\begin{aligned} \langle \mu, y_0 - \bar{y} \rangle &> \langle \mu, y_0 - (\bar{y} + y) \rangle \\ &\geq \langle \mu, y_0 \rangle, \end{aligned}$$

and thus $\langle \mu, \bar{y} \rangle < 0$ which contradicts (2.12) (recall that $\bar{y} \in T_Q(y_0)$).

Now let $\bar{y} \notin T_Q(y_0)$. Since $T_Q(y_0)$ is closed, by the second separation theorem [3, Theorem 2.14] we then obtain the existence of $\mu \in Y^*$ such that (2.12) holds and $\langle \mu, \bar{y} \rangle < 0$.

Consider arbitrary sequences $\{q^k\} \subset Q$, $\{\eta^k\} \subset Y$ and a sequence of real numbers $\{t_k\}$ such that $t_k \geq 0\ \forall\, k$ and $\{q^k - t_k \eta^k\} \to y_0$. Hence

$$\langle \mu, q^k - y_0 \rangle + t_k(-\langle \mu, \eta^k \rangle) = \langle \mu, q^k - t_k \eta^k - y_0 \rangle \to 0.$$

Note that $q^k - y_0 \in R_Q(y_0) \subset T_Q(y_0)$, and (2.12) implies that the first term in the left-hand side is nonnegative $\forall\, k$. Furthermore, inequality $\langle \mu, \bar{y} \rangle < 0$ implies that the second term in the left-hand side is nonnegative as well for all k large enough, and hence, $t_k \to 0$. The letter implies that $\{q^k\} \to y_0$. Thus, $q^k - t_k \eta^k \in Q \cap B_{\delta_1}(y_0) - \operatorname{cone} B_{\delta_2}(\bar{y})$ for all k large enough. This proves the needed inclusion (2.9) with sufficiently small $\varepsilon > 0$ and $\delta > 0$. ∎

We are now ready to prove the main result of this section.

Theorem 2.1 *Let Q be closed and convex, and let $x_0 \in D(\sigma_0)$. Let F be continuous at (σ_0, x_0) and Fréchet-differentiable with respect to x near (σ_0, x_0), and let its derivative with respect to x be continuous at (σ_0, x_0).*

If the mapping $F(\sigma_0, \cdot)$ is regular at x_0 in a direction $\bar{y} \in Y$ then there exist a neighborhood U of σ_0 and $\varepsilon > 0$, $\delta > 0$ and $c > 0$ such that the estimate (1.3) *holds for all $(\sigma, x) \in U \times B_\varepsilon(x_0)$ satisfying the inclusion*

$$F(\sigma, x) \in Q - \operatorname{cone} B_\delta(\bar{y}). \tag{2.13}$$

Proof. From the equivalent form (2.1) of the directional regularity condition it evidently follows that there exists $\bar{\eta} \in \operatorname{cone}\{\bar{y}\}$ such that

$$\bar{\eta} \in \operatorname{int}\left(F(\sigma_0,\, x_0) + \operatorname{im} \frac{\partial F}{\partial x}(\sigma_0,\, x_0) - Q\right). \tag{2.14}$$

Note that if $\bar{y} = 0$ then necessarily $\bar{\eta} = 0$.

Define the multifunction $\bar{\mathcal{F}} : X \to 2^Y$,

$$\bar{\mathcal{F}}(\xi) = F(\sigma_0,\, x_0) + \frac{\partial F}{\partial x}(\sigma_0,\, x_0)\xi - Q.$$

According to (2.14), there exists $\bar{\xi} \in X$ such that $\bar{\mathcal{F}}(\bar{\xi}) = \bar{\eta}$, and moreover, by the Robinson-Ursescu stability theorem [23, 20] (see also [3, Theorem 2.83]) it follows that the multifunction $\bar{\mathcal{F}}$ is metrically regular at $(\bar{\xi},\, \bar{\eta})$.

Fix $\bar{\varepsilon} > 0$. For each mapping $G : X \to Y$, define the multifunction $\mathcal{F}_G : X \to 2^Y$,

$$\mathcal{F}_G(\xi) = F(\sigma_0,\, x_0) + G(\xi) - Q.$$

Note that $\mathcal{F}_{\frac{\partial F}{\partial x}(\sigma_0,\, x_0)} = \bar{\mathcal{F}}$, and hence, by [3, Theorem 2.84] it follows that there exist $\bar{l} > 0$, $\delta > 0$ and $\bar{c} > 0$ such that the estimate

$$\begin{gathered}\operatorname{dist}(\bar{\xi},\, \mathcal{F}_G^{-1}(y)) \le \bar{c}\operatorname{dist}(G(\bar{\xi}) - y,\, Q - F(\sigma_0,\, x_0)) \\ \forall\, y \in B_\delta\left(\bar{\eta} - \frac{\partial F}{\partial x}(\sigma_0,\, x_0)\bar{\xi} + G(\bar{\xi})\right)\end{gathered} \tag{2.15}$$

holds for each G such that the difference mapping $G(\cdot) - \frac{\partial F}{\partial x}(\sigma_0,\, x_0)$ is Lipschitz-continuous on $B_{\bar{\varepsilon}}(\bar{\xi})$ with modulus $l \in (0,\, \bar{l})$.

It can be easily seen that there exists $\tilde{\delta}_2 \in (0,\, \delta/4]$ possessing the following property: if $\eta \in \operatorname{cone} B_{\tilde{\delta}_2}(\bar{\eta}) \setminus \{0\}$ then $\|\|\bar{\eta}\|\eta/\|\eta\| - \bar{\eta}\| \le \delta/4$. Put

$$\gamma = \begin{cases} \|\bar{\eta}\| & \text{if } \bar{\eta} \ne 0, \\ \frac{\delta}{4} & \text{if } \bar{\eta} = 0. \end{cases} \tag{2.16}$$

Set $\delta_1 = \min\{\delta/16,\, \gamma/4\}$, $\delta_2 = \|\bar{y}\|\tilde{\delta}_2/\|\bar{\eta}\|$ if $\bar{\eta} \ne 0$ (so that $\operatorname{cone} B_{\delta_2}(\bar{y}) = \operatorname{cone} B_{\tilde{\delta}_2}(\bar{\eta})$; if $\bar{\eta} = 0$, $\delta_2 > 0$ can be taken arbitrarily). Fix $(\sigma,\, x) \in \Sigma \times X$ satisfying

$$F(\sigma,\, x) \in Q \cap B_{\delta_1}(F(\sigma_0,\, x_0)) - \operatorname{cone} B_{\delta_2}(\bar{y}) \tag{2.17}$$

and such that $F(\sigma,\, x) \notin Q$ (otherwise estimate (1.3) holds trivially). Then there exists $q = q(\sigma,\, x) \in Q \cap B_{\delta_1}(F(\sigma_0,\, x_0))$ such that

$$-(F(\sigma,\, x) - q) \in \operatorname{cone} B_{\tilde{\delta}_2}(\bar{\eta}),$$

and hence, according to (2.16), and to the choice of $\tilde{\delta}_2$, it holds that

$$\left\| \frac{\gamma}{\|F(\sigma,\, x) - q\|}(F(\sigma,\, x) - q) + \bar{\eta} \right\| \le \frac{\delta}{4} \tag{2.18}$$

(note that $\|F(\sigma, x) - q\|$ cannot be equal to 0 since $F(\sigma, x) \notin Q$).

Set

$$t = t(\sigma, x, q) = \min\left\{\frac{16\,\mathrm{dist}(F(\sigma, x), Q)}{\delta}, \frac{\|F(\sigma, x) - q\|}{\gamma}\right\}. \tag{2.19}$$

Note that $t > 0$ but t tends to 0 as (σ, x) tends to (σ_0, x_0). Define the mapping $G : X \to Y$,

$$G(\xi) = G(\sigma, x; \xi) = \frac{1}{t}(F(\sigma, x + t\xi) - F(\sigma, x)), \tag{2.20}$$

and the difference mapping $\Phi : X \to Y$,

$$\Phi(\xi) = \Phi(\sigma, x; \xi) = G(\xi) - \frac{\partial F}{\partial x}(\sigma_0, x_0)\xi = \frac{1}{t}\left(F(\sigma, x + t\xi) - F(\sigma, x) - \frac{\partial F}{\partial x}(\sigma_0, x_0)t\xi\right). \tag{2.21}$$

By the mean value theorem we obtain that for (σ, x) close enough to (σ_0, x_0), and for each $\xi^1, \xi^2 \in X$

$$\|\Phi(\xi^1) - \Phi(\xi^2)\| \le \sup_{\theta \in [0, 1]} \left\|\frac{\partial F}{\partial x}(\sigma, x + t(\theta\xi^1 + (1 - \theta)\xi^2)) - \frac{\partial F}{\partial x}(\sigma_0, x_0)\right\| \|\xi^1 - \xi^2\|,$$

and hence, there exist a neighborhood U of σ_0 and $\varepsilon > 0$ such that Φ is Lipschitz-continuous on $B_{\bar\varepsilon}(\bar\xi)$ with modulus $l \in (0, \bar l)$ provided $(\sigma, x) \in U \times B_\varepsilon(x_0)$. Throughout the rest of the proof we suppose that the latter inclusion holds. Then by choosing another ("smaller") U and by reducing $\varepsilon > 0$ (if necessary), we obtain

$$\|\Phi(\bar\xi)\| \le \sup_{\theta \in [0, 1]} \left\|\frac{\partial F}{\partial x}(\sigma, x + t\theta\bar\xi) - \frac{\partial F}{\partial x}(\sigma_0, x_0)\right\| \|\bar\xi\| \le \frac{\delta}{2}. \tag{2.22}$$

Set

$$\theta = \theta(\sigma, x, q) = \frac{2\|F(\sigma, x) - q\|}{\gamma}, \tag{2.23}$$

$$\tilde y = \tilde y(\sigma, x, q) = \theta F(\sigma_0, x_0) + (1 - \theta)q. \tag{2.24}$$

Note that, by the definition of δ_1, $\theta \in (0, 1]$ provided U and $\varepsilon > 0$ are chosen appropriately. Choose an element $p = p(\sigma, x) \in Q$ such that

$$\|F(\sigma, x) - p\| \le 2\,\mathrm{dist}(F(\sigma, x), Q), \tag{2.25}$$

and set

$$\tau = \tau(\sigma, x, q) = \frac{\gamma t}{\|F(\sigma, x) - q\|} \tag{2.26}$$

$$y = y(\sigma, x, p, q) = -\frac{1}{t}(\tau(F(\sigma, x) - \tilde y) + (1 - \tau)(F(\sigma, x) - p)). \tag{2.27}$$

Note that $\tau \in (0, 1]$, and moreover, $\tau = 1$ provided $\|F(\sigma, x) - q\|/\gamma \le 16\,\mathrm{dist}(F(\sigma, x), Q)/\delta$, that is, when $t = \|F(\sigma, x) - q\|/\gamma$ (see (2.19)). Taking this into account, by (2.18), (2.19), (2.23)–(2.27), and by the definition of δ_1, we derive that

$$\begin{aligned}
\|y-\bar\eta\| &= \left\|\frac{\tau}{t}(F(\sigma,\,x)-\tilde y)+\frac{1-\tau}{t}(F(\sigma,\,x)-p)+\bar\eta\right\| \\
&\le \left\|\frac{\gamma}{\|F(\sigma,\,x)-q\|}(F(\sigma,\,x)-\tilde y)+\bar\eta\right\|+(1-\tau)\frac{\|F(\sigma,\,x)-p\|}{t} \\
&\le \left\|\frac{\gamma}{\|F(\sigma,\,x)-q\|}(F(\sigma,\,x)-q)+\bar\eta\right\|+\theta\frac{\gamma\|q-F(\sigma_0,\,x_0)\|}{\|F(\sigma,\,x)-q\|}+\frac{2\delta\,\mathrm{dist}(F(\sigma,\,x),\,Q)}{16\,\mathrm{dist}(F(\sigma,\,x),\,Q)} \\
&\le \frac{\delta}{4}+\frac{\delta}{8}+\frac{\delta}{8} \\
&= \frac{\delta}{2}.
\end{aligned}$$

Thus, by the second equality in (2.21), and by (2.22), it holds that

$$\begin{aligned}
\left\|y-\bar\eta+\frac{\partial F}{\partial x}(\sigma_0,\,x_0)\bar\xi-G(\bar\xi)\right\| &\le \|y-\bar\eta\|+\|\Phi(\bar\xi)\| \\
&\le \frac{\delta}{2}+\frac{\delta}{2} \\
&\le \delta. \qquad (2.28)
\end{aligned}$$

Hence, the estimate (2.15) must be valid for y defined in (2.27) and for G defined in (2.20) provided U and $\varepsilon>0$ are chosen appropriately. This means that there exist $\xi=\xi(\sigma,\,x,\,p,\,q)\in X$ and $\eta=\eta(\sigma,\,x,\,p,\,q)\in Q$ such that

$$G(\xi)=y+\eta-F(\sigma_0,\,x_0) \qquad (2.29)$$

and

$$\begin{aligned}
\|\xi\| &\le \|\bar\xi\|+\|\xi-\bar\xi\| \\
&\le \|\bar\xi\|+\bar c\,\mathrm{dist}(G(\bar\xi)-y,\,Q-F(\sigma_0,\,x_0)) \\
&\le \|\bar\xi\|+\bar c\|G(\bar\xi)-y\| \\
&\le \|\bar\xi\|+\bar c\left(\left\|\frac{\partial F}{\partial x}(\sigma_0,\,x_0)\bar\xi-\bar\eta\right\|+\delta\right), \qquad (2.30)
\end{aligned}$$

where (2.28) and the inclusion $0\in Q-F(\sigma_0,\,x_0)$ were taken into account. Note that the right-hand side of the last relation is a constant independent of σ, x, p and q.

Employing (2.21), (2.24), (2.27) and (2.29), we have

$$\begin{aligned}
F(\sigma,\,x+t\xi) &= t\Phi(\xi)+F(\sigma,\,x)+t\frac{\partial F}{\partial x}(\sigma_0,\,x_0)\xi \\
&= t\Phi(\xi)+t\frac{\partial F}{\partial x}(\sigma_0,\,x_0)\xi \\
&\quad +\tau(F(\sigma,\,x)-\tilde y)+\tau\tilde y+(1-\tau)(F(\sigma,\,x)-p)+(1-\tau)p \\
&= tG(\xi)-ty+\tau\tilde y+(1-\tau)p \\
&= t(\eta-F(\sigma_0,\,x_0))+\tau\tilde y+(1-\tau)p \\
&= t\eta-tF(\sigma_0,\,x_0)+\tau\theta F(\sigma_0,\,x_0)+\tau(1-\theta)q+(1-\tau)p \\
&= t\eta+(\tau\theta-t)F(\sigma_0,\,x_0)+\tau(1-\theta)q+(1-\tau)p,
\end{aligned}$$

where the right-hand side is a convex combination of η, $F(\sigma_0,\, x_0)$, p and q provided U and $\varepsilon > 0$ are chosen appropriately. However, all the elements η, $F(\sigma_0,\, x_0)$, p and q belong to the convex set Q. Hence,

$$F(\sigma,\, x + t\xi) \in Q,$$

and moreover, by (2.19) and (2.30),

$$t\|\xi\| \leq c\,\mathrm{dist}(F(\sigma,\, x),\, Q),$$

where $c = 16(\|\bar{\xi}\| + \bar{c}(\|\frac{\partial F}{\partial x}(\sigma_0,\, x_0)\bar{\xi} - \bar{\eta}\| + \delta))/\delta$.

We thus proved that (1.3) holds for all $(\sigma,\, x) \in U \times B_\varepsilon(x_0)$ satisfying (2.17). In order to completes the proof it suffices to refer to Proposition 2.2. ■

3 Directional Metric Regularity

Let $(X,\, \rho)$ be a complete metric space, and Y be a normed linear space. As will be explained below, the following definition is motivated by Theorem 2.1.

Definition 3.1 The multifunction $\Psi\colon X \to 2^Y$ is *metrically regular at a point* $(x_0,\, y_0) \in \mathrm{graph}\,\Psi$ *in a direction* $\bar{y} \in Y$, at a rate $c > 0$, if there exist $\varepsilon > 0$ and $\delta > 0$ such that the estimate

$$\mathrm{dist}(x,\, \Psi^{-1}(y)) \leq c\,\mathrm{dist}(y,\, \Psi(x)) \tag{3.1}$$

holds for all $(x,\, y) \in B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ satisfying the inclusion

$$y \in \Psi(x) + \mathrm{cone}\, B_\delta(\bar{y}). \tag{3.2}$$

Evidently, metric regularity in a direction $\bar{y} = 0$ is equivalent to the usual metric regularity. Moreover, if the latter holds, directional metric regularity holds in any direction $\bar{y} \in Y$, including $\bar{y} = 0$. At the same time, directional metric regularity can hold when the usual metric regularity is violated; see Example 3.1 below.

Recall that the multifunction $\Psi : X \to 2^Y$ is said to be *lower* (or *inner*) *semicontinuous at a point* $(x_0,\, y_0) \in \mathrm{graph}\,\Psi$ if for any sequence $\{x^k\} \subset X$ convergent to x_0 there exists a sequence $\{y^k\} \subset Y$ convergent to y_0 such that $y^k \in \Psi(x^k)\ \forall\, k$ (see, e.g., [16, Definition 1.63]).

The next theorem follows the pattern of [3, Theorem 2.84], [16, Theorem 4.25]; it says that the property of directional metric regularity of Ψ in a given direction $\bar{y}$ is stable subject to small Lipschitzian single-valued perturbations of Ψ. For the usual notion of metric regularity, this property was studied, e.g., in [9, 7, 8]. Yet another reference to be mentioned in relation with this property is [6], where the importance of stability of regularity properties with respect to perturbations is already completely clear.

Theorem 3.1 *Let* $(x_0,\, y_0) \in \mathrm{graph}\,\Psi$. *Assume that the multifunction* Ψ *is closed, lower semicontinuous at* $(x_0,\, y_0)$ *and metrically regular at* $(x_0,\, y_0)$ *in a direction* $\bar{y} \in Y$, *at a rate* $c > 0$. *Let* $\varepsilon > 0$ *and* $\delta > 0$ *be chosen according to Definition 3.1, and set*

$$\alpha = \begin{cases} \frac{\delta}{\|\bar{y}\|} & \text{if } \|\bar{y}\| \geq \delta, \\ +\infty & \text{if } \|\bar{y}\| < \delta, \end{cases} \qquad \beta(\alpha) = \begin{cases} \frac{2}{1+\alpha} & \text{if } \alpha < +\infty, \\ 0 & \text{if } \alpha = +\infty. \end{cases} \tag{3.3}$$

Then for any mapping $\Phi : X \to Y$ which is Lipschitz-continuous on $B_\varepsilon(x_0)$ with modulus $l > 0$ such that

$$c\,l < \min\{1,\, \alpha/5\}, \tag{3.4}$$

the multifunction $\Psi + \Phi$ is metrically regular at $(x_0,\, y_0 + \Phi(x_0))$ in the direction $\bar{y}$, at a rate $\tilde{c} = c(1 - c\,l)^{-1}(1 + \beta(\alpha))$.

In (3.3), the possibility of $\alpha = +\infty$ is needed only in order to cover the case of usual metric regularity corresponding to $\bar{y} = 0$. In the latter case, Theorem 3.1 reduces to [16, Theorem 4.25] but with an extraneous assumption of lower semicontinuity. It is possible that this assumption can actually be removed in Theorem 3.1, though the authors did not manage to avoid it.

Remark 3.1 As can be seen from the proof below, the assertion of Theorem 3.1 can be replaced by a somewhat stronger one: Under the assumptions of this theorem, for each $l > 0$ satisfying (3.4) there exist $\tilde{\varepsilon} > 0$ and $\tilde{\delta} > 0$ such that the estimate

$$\operatorname{dist}(x,\, (\Psi + \Phi)^{-1}(y)) \le \tilde{c}\operatorname{dist}(y,\, \Psi(x) + \Phi(x)) \tag{3.5}$$

holds for any mapping $\Phi : X \to Y$ which is Lipschitz-continuous on $B_\varepsilon(x_0)$ with modulus l, and for all $(x,\, y) \in B_{\tilde{\varepsilon}}(x_0) \times B_{\tilde{\varepsilon}}(y_0 + \Phi(x_0))$ satisfying the inclusion

$$y \in \Psi(x) + \Phi(x) + \operatorname{cone} B_{\tilde{\delta}}(\bar{y}). \tag{3.6}$$

That is, $\tilde{\varepsilon}$ and $\tilde{\delta}$ do not depend on a specific Φ but only on ε, δ, c, $\|\bar{y}\|$ and l.

Proof. Let $\varepsilon > 0$ and $\delta > 0$ be chosen according to Definition 3.1. Fix arbitrary $\tilde{\varepsilon} \in (0,\, \varepsilon]$, $\tilde{\delta} \in (0,\, \delta)$ and $\hat{\varepsilon} > 0$ satisfying the following set of conditions:

$$\tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1 - \gamma(\hat{\varepsilon}))^{-1}(1 + \beta(\alpha))(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) < \varepsilon, \tag{3.7}$$

$$\tilde{\varepsilon} + l\tilde{\varepsilon} + \left(\gamma(\hat{\varepsilon})(1 - \gamma(\hat{\varepsilon}))^{-1}(1 + \beta(\alpha)) + \frac{\beta(\alpha)}{2}\right)(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) < \varepsilon, \tag{3.8}$$

$$\gamma(\hat{\varepsilon}) < \min\left\{1,\, \frac{\alpha(\delta - \tilde{\delta})}{5\delta}\right\}, \tag{3.9}$$

where $\omega(\tilde{\varepsilon}) = \sup_{x \in B_{\tilde{\varepsilon}}(x_0)} \operatorname{dist}(y_0,\, \Psi(x))$, $\gamma(\hat{\varepsilon}) = c\,l(1 + \hat{\varepsilon})$ (note that $\omega(\tilde{\varepsilon}) \to 0$ as $\tilde{\varepsilon} \to 0$ because of the lower semicontinuity Ψ at $(x_0,\, y_0)$, and recall (3.4)).

Let $(x,\, y) \in B_{\tilde{\varepsilon}}(x_0) \times B_{\tilde{\varepsilon}}(y_0 + \Phi(x_0))$ satisfying (3.6) be fixed. In order to prove estimate (3.5) it suffices to establish the existence of $\chi(x) \in (\Psi + \Phi)^{-1}(y)$ such that

$$\rho(x,\, \chi(x)) \le c(1 - c\,l)^{-1}(1 + \beta(\alpha))\operatorname{dist}(y,\, \Psi(x) + \Phi(x)). \tag{3.10}$$

The needed point $\chi(x)$ will be defined by means of the auxiliary iterative process. For that purpose set $t = t(x, y) = \mathrm{dist}(y, \Psi(x) + \Phi(x))$ and define the sequence $\{\tau_k\} \subset \mathbf{R}_+$ by setting

$$\tau_1 = \begin{cases} \frac{\delta - \tilde{\delta}}{(\|\bar{y}\| + \delta)\delta} t & \text{if } \|\bar{y}\| \geq \delta, \\ 0 & \text{if } \|\bar{y}\| < \delta, \end{cases} \qquad \tau_{k+1} = \frac{2}{5}\tau_k, \ k = 1, 2, \ldots.$$

According to (3.3),

$$\tau_1 \|\bar{y}\| \leq \frac{\beta(\alpha)}{2} t. \tag{3.11}$$

Note that by the definition of $\omega(\tilde{\varepsilon})$

$$\begin{aligned} t &= \mathrm{dist}(y - \Phi(x_0) + \Phi(x_0) - \Phi(x), \Psi(x)) \\ &\leq \|y - y_0 - \Phi(x_0)\| + \|\Phi(x) - \Phi(x_0)\| + \mathrm{dist}(y_0, \Psi(x)) \\ &\leq \tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon}). \end{aligned} \tag{3.12}$$

We shall construct a sequence $\{x^k\} \subset X$ such that $x^1 = x$ and $\forall\, k = 1, 2, \ldots$

$$y - \Phi(x^k) - \tau_k \bar{y} \in \Psi(x^{k+1}), \tag{3.13}$$

$$\rho(x^{k+2}, x^{k+1}) \leq \gamma(\hat{\varepsilon})\rho(x^{k+1}, x^k) + \frac{\gamma(\hat{\varepsilon})}{l}(\tau_k - \tau_{k+1})\|\bar{y}\|, \tag{3.14}$$

$$\rho(x^{k+1}, x_0) \leq \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1 - \gamma(\hat{\varepsilon}))^{-1}(1 + \beta(\alpha))(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})), \tag{3.15}$$

and if $\|\bar{y}\| \geq \delta$ then

$$\|\Phi(x^{k+1}) - \Phi(x^k)\| \leq \delta(\tau_k - \tau_{k+1}). \tag{3.16}$$

By (3.6) we obtain the existence of $\theta \geq 0$ and $\eta \in Y$ such that $\|\eta\| \leq \tilde{\delta}$ and

$$y - \Phi(x^1) - \theta(\bar{y} + \eta) \in \Psi(x^1). \tag{3.17}$$

Note that if $\theta = 0$ then $x = x^1 \in (\Psi + \Phi)^{-1}(y)$, and we are done. Thus, let $\theta > 0$.

Set

$$y^1 = y - \Phi(x^1) - \tau_1 \bar{y}, \quad \eta_0 = \theta(\bar{y} + \eta), \quad \eta^1 = \eta_0 - \tau_1 \bar{y}.$$

By (3.8), (3.11) and (3.12) we then derive

$$\begin{aligned} \|y^1 - y_0\| &= \|y^1 + \Phi(x_0) - y_0 - \Phi(x_0)\| \\ &\leq \|y - y_0 - \Phi(x_0)\| + \|\Phi(x_0) - \Phi(x^1)\| + \tau_1\|\bar{y}\| \\ &\leq \tilde{\varepsilon} + l\tilde{\varepsilon} + \frac{\beta(\alpha)}{2}(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) \\ &< \varepsilon. \end{aligned} \tag{3.18}$$

Furthermore,

$$\begin{aligned} \|\eta_0\| &\leq \theta(\|\bar{y}\| + \tilde{\delta}) \\ &< \theta(\|\bar{y}\| + \delta), \end{aligned}$$

and hence,

$$\begin{aligned}\theta &> \frac{\|\eta_0\|}{\|\bar{y}\|+\delta} \\ &\geq \frac{t}{\|\bar{y}\|+\delta} \\ &\geq \tau_1, \end{aligned} \tag{3.19}$$

where it was taken into account that, by (3.17), $y - \eta_0 \in \Psi(x^1) + \Phi(x^1)$, and hence, by the definition of t, it holds that $t \leq \|\eta_0\|$. From (3.19) (including the intermediate inequalities) and the definition of τ_1 it follows that

$$\begin{aligned}\frac{\theta\|\eta\|}{\theta-\tau_1} &\leq \frac{\theta\tilde{\delta}}{\theta - \frac{\delta-\tilde{\delta}}{(\|\bar{y}\|+\delta)\delta}t} \\ &\leq \frac{\theta\tilde{\delta}}{\theta - \theta\frac{\delta-\tilde{\delta}}{\delta}} \\ &= \delta,\end{aligned}$$

and hence, by (3.17),

$$\begin{aligned}\eta^1 &= (\theta-\tau_1)\bar{y} + \theta\eta \\ &= (\theta-\tau_1)\left(\bar{y} + \frac{\theta\eta}{\theta-\tau_1}\right) \\ &\in \operatorname{cone} B_\delta(\bar{y}).\end{aligned} \tag{3.20}$$

Taking into account the equality $y^1 = y - \Phi(x^1) - \theta(\bar{y}+\eta) + \eta^1$, we conclude by (3.17) and (3.20) that

$$y^1 \in \Psi(x^1) + \operatorname{cone} B_\delta(\bar{y}),$$

that is, (3.2) holds with $(x,\, y)$ replaced by $(x^1,\, y^1) \in B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ (see (3.18)). Thus, by metric regularity of Ψ at a point $(x_0,\, y_0)$ in a direction $\bar{y}$, there exists $x^2 \in X$ such that

$$y - \Phi(x^1) - \tau_1\bar{y} \in \Psi(x^2), \tag{3.21}$$

$$\begin{aligned}\rho(x^2,\, x^1) &\leq c(1+\hat{\varepsilon})\operatorname{dist}(y - \Phi(x^1) - \tau_1\bar{y},\, \Psi(x^1)) \\ &\leq c(1+\hat{\varepsilon})(t + \tau_1\|\bar{y}\|) \\ &\leq \frac{\gamma(\hat{\varepsilon})}{l}\left(1 + \frac{\beta(\alpha)}{2}\right)t,\end{aligned} \tag{3.22}$$

where the definition of t and (3.11) were taken into account. In particular, (3.13) holds for $k = 1$.

Employing (3.7), (3.12), (3.22) we derive

$$\rho(x^2,\, x_0) \leq \rho(x^1,\, x_0) + \rho(x^2,\, x^1)$$

$$
\begin{aligned}
&\leq \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}\left(1 + \frac{\beta(\alpha)}{2}\right)t \\
&\leq \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}\left(1 + \frac{\beta(\alpha)}{2}\right)(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) \\
&< \varepsilon,
\end{aligned} \tag{3.23}
$$

and in particular, (3.15) holds for $k = 1$.

Furthermore, if $\|\bar{y}\| \geq \delta$ then by (3.3), (3.9), (3.22) and (3.23) we derive

$$
\begin{aligned}
\|\Phi(x^2) - \Phi(x^1)\| &\leq \gamma(\hat{\varepsilon})\left(1 + \frac{\beta(\alpha)}{2}\right)t \\
&< \frac{\alpha(\delta - \tilde{\delta})}{5\delta}\left(1 + \frac{1}{1+\alpha}\right)t \\
&= \frac{\alpha(\delta - \tilde{\delta})(2+\alpha)}{5(1+\alpha)\delta}t \\
&\leq \frac{3\alpha(\delta - \tilde{\delta})}{5(1+\alpha)\delta}t \\
&= \delta(\tau_1 - \tau_2),
\end{aligned} \tag{3.24}
$$

that is, (3.16) holds for $k = 1$.

Set

$$
q^2 = y - \Phi(x^1) - \tau_1\bar{y},
$$
$$
y^2 = y - \Phi(x^2) - \tau_2\bar{y}, \quad \eta^2 = (\tau_1 - \tau_2)\bar{y} + \Phi(x^1) - \Phi(x^2).
$$

Note that $q^2 \in \Psi(x^2)$ by (3.21), and by (3.24) we conclude that $\eta^2 \in \operatorname{cone} B_\delta(\bar{y})$ ((3.24) holds only if $\|\bar{y}\| \geq \delta$, but otherwise, $\operatorname{cone} B_\delta(\bar{y}) = Y$).

By (3.8), (3.11) and (3.12), and by (3.15) (for $k = 1$) it follows that

$$
\begin{aligned}
\|y^2 - y_0\| &= \|y - y_0 - \Phi(x_0)\| + \|\Phi(x^2) - \Phi(x_0)\| + \tau_2\|\bar{y}\| \\
&\leq \tilde{\varepsilon} + l\tilde{\varepsilon} + \gamma(\hat{\varepsilon})(1 - \gamma(\hat{\varepsilon}))^{-1}(1 + \beta(\alpha))(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon}))) + \frac{\beta(\alpha)}{2}(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) \\
&< \varepsilon.
\end{aligned} \tag{3.25}
$$

The inclusions $q^2 \in \Psi(x^2)$ and $\eta^2 \in \operatorname{cone} B_\delta(\bar{y})$ imply that

$$
\begin{aligned}
y^2 &= y - \Phi(x^1) - \tau_1\bar{y} + \Phi(x^1) + \tau_1\bar{y} - \Phi(x^2) - \tau_2\bar{y} \\
&= q^2 + \eta^2 \\
&\in \Psi(x^2) + \operatorname{cone} B_\delta(\bar{y}),
\end{aligned}
$$

that is, (3.2) holds with (x, y) replaced by $(x^2, y^2) \in B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ (see (3.23), (3.25)). Thus, by metric regularity of Ψ at a point (x_0, y_0) in a direction $\bar{y}$, there exists $x^3 \in X$ such that

$$
y - \Phi(x^2) - \tau_2\bar{y} \in \Psi(x^3), \tag{3.26}
$$

$$
\begin{aligned}
\rho(x^3, x^2) &\leq c(1+\hat{\varepsilon})\,\mathrm{dist}(y - \Phi(x^2) - \tau_2\bar{y},\, \Psi(x^2)) \\
&\leq c(1+\hat{\varepsilon})\|y - \Phi(x^2) - \tau_2\bar{y} - q^2\| \\
&\leq c(1+\hat{\varepsilon})(\|\Phi(x^2) - \Phi(x^1)\| + (\tau_1 - \tau_2)\|\bar{y}\|) \\
&\leq \gamma(\hat{\varepsilon})\rho(x^2, x^1) + \frac{\gamma(\hat{\varepsilon})}{l}(\tau_1 - \tau_2)\|\bar{y}\|,
\end{aligned} \tag{3.27}
$$

where the definition of q^2 was taken into account. In particular, (3.13) holds for $k = 2$, and (3.14) holds for $k = 1$.

Employing (3.11), (3.22), (3.27) we derive

$$
\begin{aligned}
\rho(x^3, x^1) &\leq \rho(x^2, x^1) + \rho(x^3, x^2) \\
&\leq (1+\gamma(\hat{\varepsilon}))\rho(x^2, x^1) + \frac{\gamma(\hat{\varepsilon})}{l}(\tau_1 - \tau_2)\|\bar{y}\| \\
&\leq (1+\gamma(\hat{\varepsilon}))\left(\rho(x^2, x^1) + \frac{\gamma(\hat{\varepsilon})}{l}\tau_1\|\bar{y}\|\right) \\
&\leq \frac{\gamma(\hat{\varepsilon})}{l}(1+\gamma(\hat{\varepsilon}))\left(\left(1 + \frac{\beta(\alpha)}{2}\right)t + \tau_1\|\bar{y}\|\right) \\
&\leq \frac{\gamma(\hat{\varepsilon})}{l}(1+\gamma(\hat{\varepsilon}))(1+\beta(\alpha))t,
\end{aligned}
$$

and thus, by (3.7) and (3.12),

$$
\begin{aligned}
\rho(x^3, x_0) &\leq \rho(x^1, x_0) + \rho(x^3, x^1) \\
&\leq \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1+\gamma(\hat{\varepsilon}))(1+\beta(\alpha))t \\
&\leq \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1-\gamma(\hat{\varepsilon}))^{-1}(1+\beta(\alpha))(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) \\
&< \varepsilon,
\end{aligned} \tag{3.28}
$$

where the evident inequality $1 + \gamma(\hat{\varepsilon}) < (1 - \gamma(\hat{\varepsilon}))^{-1}$ was also employed. In particular, (3.15) holds for $k = 2$.

Furthermore, if $\|\bar{y}\| \geq \delta$ then by (3.3), (3.9), (3.24) and (3.28), and by the intermediate inequalities in (3.27) we derive

$$
\begin{aligned}
\|\Phi(x^3) - \Phi(x^2)\| &\leq l\rho(x^3, x^2) \\
&\leq \gamma(\hat{\varepsilon})(\|\Phi(x^2) - \Phi(x^1)\| + (\tau_2 - \tau_1)\|\bar{y}\|) \\
&\leq \gamma(\hat{\varepsilon})(\delta + \|\bar{y}\|)(\tau_1 - \tau_2) \\
&< \frac{\alpha}{5}(\delta + \|\bar{y}\|)\frac{5}{2}\left(1 - \frac{2}{5}\right)\tau_2 \\
&= \frac{\delta}{2\|\bar{y}\|}(\delta + \|\bar{y}\|)(\tau_2 - \tau_3) \\
&\leq \delta(\tau_2 - \tau_3),
\end{aligned}
$$

that is, (3.16) holds for $k = 2$.

Suppose now that for some $s \geq 3$ we have already constructed points $x^k \in X$, $k = 1, \ldots, s$, such that (3.13), (3.15), and (3.16) if $\|\bar{y}\| \geq \delta$, hold for each $k = 1, \ldots, s-1$, and (3.14) holds for each $k = 1, \ldots, s-2$. Set

$$q^s = y - \Phi(x^{s-1}) - \tau_{s-1}\bar{y},$$

$$y^s = y - \Phi(x^s) - \tau_s\bar{y}, \quad \eta^s = (\tau_{s-1} - \tau_s)\bar{y} + \Phi(x^{s-1}) - \Phi(x^s).$$

Note that $q^s \in \Phi(x^s)$ by (3.13) (with $k = s-1$), and by (3.16) (with $k = s-1$) we conclude that $\eta^s \in \operatorname{cone} B_\delta(\bar{y})$ ((3.16) holds only if $\|\bar{y}\| \geq \delta$, but otherwise, $\operatorname{cone} B_\delta(\bar{y}) = Y$).

By (3.8), (3.11) and (3.12), and by (3.15) (for $k = s-1$) it follows that

$$\begin{aligned} \|y^s - y_0\| &= \|y - y_0 - \Phi(x_0)\| + \|\Phi(x^s) - \Phi(x_0)\| + \tau_s\|\bar{y}\| \\ &\leq \tilde{\varepsilon} + l\tilde{\varepsilon} + \gamma(\hat{\varepsilon})(1-\gamma(\hat{\varepsilon}))^{-1}(1+\beta(\alpha))(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon}))) + \frac{\beta(\alpha)}{2}(\tilde{\varepsilon} + l\tilde{\varepsilon} + \omega(\tilde{\varepsilon})) \\ &< \varepsilon. \end{aligned} \tag{3.29}$$

The inclusions $q^s \in \Phi(x^s)$ and $\eta^s \in \operatorname{cone} B_\delta(\bar{y})$ imply that

$$\begin{aligned} y^s &= y - \Phi(x^{s-1}) - \tau_{s-1}\bar{y} + \Phi(x^{s-1}) + \tau_{s-1}\bar{y} - \Phi(x^s) - \tau_s\bar{y} \\ &= q^s + \eta^s \\ &\in \Psi(x^s) + \operatorname{cone} B_\delta(\bar{y}), \end{aligned}$$

that is, (3.2) holds with (x, y) replaced by $(x^s, y^s) \in B_\varepsilon(x_0) \times B_\varepsilon(y_0)$ (see (3.8), (3.15), (3.29)). Thus, by metric regularity of Ψ at a point (x_0, y_0) in a direction $\bar{y}$, there exists $x^{s+1} \in X$ such that

$$y - \Phi(x^s) - \tau_s\bar{y} \in \Psi(x^{s+1}),$$

$$\begin{aligned} \rho(x^{s+1}, x^s) &\leq c(1+\hat{\varepsilon}) \operatorname{dist}(y - \Phi(x^s) - \tau_s\bar{y}, \Psi(x^s)) \\ &\leq c(1+\hat{\varepsilon})\|y - \Phi(x^s) - \tau_s\bar{y} - q^s\| \\ &\leq c(1+\hat{\varepsilon})(\|\Phi(x^s) - \Phi(x^{s-1})\| + (\tau_{s-1} - \tau_s)\|\bar{y}\|) \\ &\leq \gamma(\hat{\varepsilon})\rho(x^s, x^{s-1}) + \frac{\gamma(\hat{\varepsilon})}{l}(\tau_{s-1} - \tau_s)\|\bar{y}\|, \end{aligned} \tag{3.30}$$

where the definition of q^s was taken into account. In particular, (3.13) holds for $k = s$, and (3.14) holds for $k = s-1$.

Employing (3.14) we derive that for each $k = 1, \ldots, s-1$

$$\begin{aligned} \sum_{i=1}^{k} \rho(x^{i+2}, x^{i+1}) &\leq \sum_{i=1}^{k} \left(\gamma(\hat{\varepsilon})\rho(x^{i+1}, x^i) + \frac{\gamma(\hat{\varepsilon})}{l}(\tau_i - \tau_{i+1})\|\bar{y}\| \right) \\ &\leq \gamma(\hat{\varepsilon}) \sum_{i=1}^{k} \rho(x^{i+1}, x^i) + \frac{\gamma(\hat{\varepsilon})}{l}\tau_1\|\bar{y}\|. \end{aligned}$$

It can be easily seen by induction that the latter property implies the estimate

$$\sum_{k=1}^{s-1} \rho(x^{k+2}, x^{k+1}) \leq \gamma(\hat{\varepsilon})(1-\gamma(\hat{\varepsilon}))^{-1}(\rho(x^2, x^1) + l^{-1}\tau_1\|\bar{y}\|). \tag{3.31}$$

Hence, by (3.11), (3.22),

$$\begin{aligned}\rho(x^{s+1},\, x^1) &\le \rho(x^2,\, x^1) + \sum_{i=1}^{s-1} \rho(x^{i+2},\, x^{i+1}) \\ &\le \left(1+\gamma(\hat{\varepsilon})(1-\gamma(\hat{\varepsilon}))^{-1}\right)\rho(x^2,\, x^1) + \frac{\gamma(\hat{\varepsilon})}{l}(1-\gamma(\hat{\varepsilon}))^{-1}\tau_1\|\bar{y}\| \\ &\le \frac{\gamma(\hat{\varepsilon})}{l}(1-\gamma(\hat{\varepsilon}))^{-1}(1+\beta(\alpha))t, \qquad (3.32)\end{aligned}$$

and thus, by (3.7) and (3.12),

$$\begin{aligned}\rho(x^{s+1},\, x_0) &\le \rho(x^1,\, x_0) + \rho(x^{s+1},\, x^1) \\ &\le \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1-\gamma(\hat{\varepsilon}))^{-1}(1+\beta(\alpha))t \\ &\le \tilde{\varepsilon} + \frac{\gamma(\hat{\varepsilon})}{l}(1-\gamma(\hat{\varepsilon}))^{-1}(1+\beta(\alpha))(\tilde{\varepsilon}+l\tilde{\varepsilon}+\omega(\tilde{\varepsilon})) \\ &< \varepsilon. \qquad (3.33)\end{aligned}$$

In particular, (3.15) holds for $k = s$.

Finally, if $\|\bar{y}\| \ge \delta$ then by (3.3), (3.9), (3.16) (with $k = s-1$) and (3.33), and by the intermediate inequalities in (3.30) we derive

$$\begin{aligned}\|\Phi(x^{s+1}) - \Phi(x^s)\| &\le l\rho(x^{s+1},\, x^s) \\ &\le \gamma(\hat{\varepsilon})(\|\Phi(x^s) - \Phi(x^{s-1})\| + (\tau_{s-1}-\tau_s)\|\bar{y}\|) \\ &\le \gamma(\hat{\varepsilon})(\delta + \|\bar{y}\|)(\tau_{s-1}-\tau_s) \\ &< \frac{\alpha}{5}(\delta+\|\bar{y}\|)\frac{5}{2}\left(1-\frac{2}{5}\right)\tau_s \\ &= \frac{\delta}{2\|\bar{y}\|}(\delta+\|\bar{y}\|)(\tau_s-\tau_{s+1}) \\ &\le \delta(\tau_s-\tau_{s+1}),\end{aligned}$$

that is, (3.16) holds for $k = s$.

The sequence $\{x^k\}$ with the needed properties is thus constructed. Moreover, as was shown above, (3.14) implies that (3.31) and (3.32) hold for each $s = 2, 3, \ldots$. Clearly, (3.31) implies that $\{x^k\}$ is a Cauchy sequence, and by completeness of the metric space (X, ρ), this sequence converges to some element $\chi(x) \in B_\varepsilon(x_0)$, where the last inclusion follows from (3.7) and (3.15). Since $\tau_k \to 0$ as $k \to \infty$, by passing onto the limit in (3.13) we conclude that $\chi(x) \in (\Psi + \Phi)^{-1}(y)$, where closedness Ψ and continuity of Φ on $B_\varepsilon(x_0)$ where taken into account. Finally, since $\hat{\varepsilon} > 0$ can be taken arbitrarily small, (3.32) implies (3.10). ■

The set cone $B_\delta(\bar{y})$ in the right-hand side of (3.2) can be regarded as a *conic neighborhood* of $\bar{y}$. Note that α defined in (3.3) is invariant with respect to the choice of specific $\bar{y}$ and δ defining the same conic neighborhood, and it is natural to refer to this quantity as the *radius* of the conic neighborhood in question.

We now turn our attention to the multifunctions of the form $\Psi(x) = \Psi_F(x) = F(x) - Q$, where $F : X \to Y$ is a given mapping and $Q \subset Y$ is a given set. Note that if F is continuous at x_0 then this multifunction is automatically lower semicuntinuous at (x_0, y_0) for any $y_0 \in \Psi(x_0)$. Being applied to such multifunction, estimate (3.1) takes the form (1.4), while condition (3.2) takes the form

$$F(x) - y \in Q - \operatorname{cone} B_\delta(\bar{y}). \tag{3.34}$$

Definition 3.1 applied to $\Psi = \Psi_F$ and $y_0 = 0$ takes the following form.

Definition 3.2 The mapping $F : X \to Y$ is *metrically regular at $x_0 \in F^{-1}(Q)$ with respect to Q in a direction $\bar{y} \in Y$*, at a rate $c > 0$, if there exist $\varepsilon > 0$ and $\delta > 0$ such that the estimate (1.4) holds for all $(x, y) \in B_\varepsilon(x_0) \times B_\varepsilon(0)$ satisfying the inclusion (3.34).

$$F(x) - y \in Q - \operatorname{cone} B_\delta(\bar{y}). \tag{3.35}$$

Throughout the rest of the paper let X and Y be Banach spaces. For a nonparametric mapping F, directional regularity condition in a direction $\bar{y}$ takes the form

$$0 \in \operatorname{int}(F(x_0) + \operatorname{im} F'(x_0) - \operatorname{cone}\{\bar{y}\} - Q), \tag{3.36}$$

and if Q is closed and convex then according to Theorem 2.1 (applied to $F(\sigma, x) = F(x) - \sigma$, $\sigma = y$), under the appropriate smoothness assumptions, the latter condition implies metric regularity in a direction $\bar{y}$. The converse implication can be derived from Theorem 3.1, which results in the following

Proposition 3.1 *Let Q be closed and convex, and let $x_0 \in F^{-1}(Q)$. Let F be Fréchet-differentiable near x_0, and let its derivative be continuous at x_0.*

Then F is metrically regular at x_0 with respect to Q in a direction $\bar{y} \in Y$ if and only if it is regular at x_0 in this direction.

Proof. Let F be metrically regular at x_0 with respect to Q in a direction $\bar{y} \in Y$, at a rate $c > 0$. Define the mapping $\Phi : X \to Y$,

$$\Phi(x) = F(x_0) + F'(x_0)(x - x_0) - F(x),$$

then $F + \Phi$ is a linearization of F at x_0. By the mean value theorem, for all $x^1, x^2 \in X$ close enough to x_0 we obtain

$$\begin{aligned}\|\Phi(x^1) - \Phi(x^2)\| &= \|F(x^1) - F(x^2) - F'(x_0)(x^1 - x^2)\| \\ &\leq \sup_{\theta \in [0,1]} \|F'(\theta x^1 + (1-\theta)x^2) - F'(x_0)\| \|x^1 - x^2\|,\end{aligned}$$

and hence, Φ is Lipschitz-continuous near x_0 with modulus l, with $l > 0$ as small as needed. Applying Theorem 3.1 to $\Psi = \Psi_F$, we conclude that the linearized mapping $F+\Phi$ is metrically regular at x_0 with respect to Q in a direction $\bar{y} \in Y$, at some rate $\tilde{c} > 0$ (note that $\Psi_F + \Phi = \Psi_{F+\Phi}$). This means that there exist $\tilde{\varepsilon} > 0$ and $\tilde{\delta} > 0$ such that the estimate

$$\operatorname{dist}(x,\, x_0 + (F'(x_0))^{-1}(Q + y - F(x_0))) \leq \tilde{c}\operatorname{dist}(F(x_0) + F'(x_0)(x - x_0) - y,\, Q) \quad (3.37)$$

holds for $(x,\, y) \in B_{\tilde{\varepsilon}}(x_0) \times B_{\tilde{\varepsilon}}(0)$ satisfying the inclusion

$$F(x_0) + F'(x_0)(x - x_0) - y \in Q - \operatorname{cone} B_{\tilde{\delta}}(\bar{y}). \quad (3.38)$$

Take $x = x_0$, $y = -\theta\eta$, where $\eta \in B_{\tilde{\delta}}(\bar{y})$ and $\theta \geq 0$. Then (3.38) is evidently satisfied, and $y \in B_{\tilde{\varepsilon}}(0)$ for all $\theta > 0$ small enough (specifically, for all $\theta \in (0,\, \tilde{\varepsilon}/(\|\bar{y}\| + \tilde{\delta}))$). Hence, (3.37) holds for chosen x and y, which implies that for all $\eta \in B_{\tilde{\delta}}(\bar{y})$ and all $\theta > 0$ small enough

$$(F'(x_0))^{-1}(Q + \theta\eta - F(x_0)) \neq \emptyset,$$

and hence, there exist $\xi \in X$ and $q \in Q$ such that

$$F'(x_0)\xi = q + \theta\eta - F(x_0),$$

i.e.,

$$\theta\eta \in F(x_0) + \operatorname{im} F'(x_0) - Q.$$

It follows that

$$B_{\tilde{\delta}}(\bar{y}) \subset \operatorname{im} F'(x_0) - R_Q(F(x_0)).$$

It remains to employ Proposition 2.1 (see (2.2)). ■

As mentioned above, directional metric regularity can hold when the usual metric regularity is violated. Moreover, let, e.g., $\operatorname{int} Q \neq \emptyset$ (which in particular covers the case of finitely many inequality constraints). It can be shown that in this case directional regularity condition (1.5), and hence, directional metric regularity condition hold in any direction $\bar{y} \in -\operatorname{int} R_Q(F(\sigma_0,\, x_0)) \neq \emptyset$.

Example 3.1 Let $X = Y = \mathbf{R}^2$, $F(x) = (x_1,\, x_1^2 - x_2^2)$, $Q = \mathbf{R}^2_+$. Robinson's CQ does not hold at $x_0 = 0$, and hence, the mapping F is not metrically regular at x_0. Moreover, estimate (1.4) does not hold even on the subspaces $\{x_0\} \times Y$ and $X \times \{0\}$. Indeed, if, e.g., $y = (0,\, y_2)$ with $y_2 < 0$, it holds that $\operatorname{dist}(x_0,\, F^{-1}(Q - y)) = (-y_2)^{1/2}$, and the estimate (1.4) does not hold even for $x = x_0$. Moreover, if, e.g., $x = (0,\, x_2)$ with $x_2 \neq 0$ then $\operatorname{dist}(x,\, F^{-1}(Q)) = |x_2|/\sqrt{2}$, while $\operatorname{dist}(F(x),\, Q) = x_2^2$, and the estimate (1.4) does not hold even for $y = 0$.

At the same time, directional regularity condition (3.36) holds at x_0 in any direction $\bar{y} \in \mathbf{R}^2$ with $\bar{y}_2 < 0$, and hence, F is metrically regular at x_0 in each such direction.

To complete this section we note that Theorem 2.1 can actually be derived from a "uniform version" of Theorem 3.1, following the line of the argument in [3, pp. 63, 64], justifying Robinson's stability theorem.

4 Applications to Sensitivity Theory

Let Σ be a normed linear space. As an application of Theorem 2.1, we next show how it can be used in order to directly (that is, without employing any additional tools, with the only exception for the mean value theorem) obtain some principal lemmas playing the crucial role in sensitivity analysis under the more special directional regularity condition (1.6). We emphasize that both results presented below are known: the difference is only in the proofs. The first result is [3, Lemma 4.10].

Lemma 4.1 *Let Q be closed and convex, and let $x_0 \in D(\sigma_0)$. Let F possess the Lipschitz-continuous derivative near $(\sigma_0,\, x_0)$.*

If (1.6) holds at x_0 with respect to a direction $d \in \Sigma$, then there exist $\bar{t} > 0$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ and $a > 0$ possessing the following property: for any mappings $\rho(\cdot) : \mathbf{R}_+ \to \Sigma$ and $x(\cdot) : \mathbf{R}_+ \to X$ such that $\rho(t) = o(t)$ and the estimates

$$\|x(t) - x_0\| \leq \varepsilon_1 t^{1/2} \tag{4.1}$$

and

$$\mathrm{dist}(F(\sigma_0 + td + \rho(t),\, x(t)),\, Q) \leq \varepsilon_2 t \tag{4.2}$$

hold for all $t \geq 0$ small enough, the estimate

$$\mathrm{dist}(x(t),\, D(\sigma_0 + td + \rho(t))) \leq a\left(1 + \frac{\|x(t) - x_0\|}{t}\right)\mathrm{dist}(F(\sigma_0 + td + \rho(t),\, x(t)),\, Q) \tag{4.3}$$

holds $\forall\, t \in (0,\, \bar{t}]$.

Proof. As was already mentioned in Section 1, (1.6) precisely coincides with (1.5) with $\bar{y} = -\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d$. For this $\bar{y}$, define $\varepsilon > 0$, $\delta > 0$ and $c > 0$ according to Theorem 2.1. Let $l > 0$ stand for the Lipschitz constant of F, and $L > 0$ stand for the Lipschitz constant for the derivative of F on $B_\varepsilon(\sigma_0) \times B_\varepsilon(x_0)$ (ε can be reduced, if necessary). For each $t > 0$ put $\sigma(t) = \sigma_0 + td + \rho(t)$. Set $\varepsilon_1 = (\delta/6L)^{1/2}$, $\varepsilon_2 = \delta/12$, and choose $\bar{t} > 0$ such that $\forall\, t \in (0,\, \bar{t}]$

$$\delta t^{1/2} \leq \varepsilon, \quad \|td + \rho(t)\| \leq \varepsilon, \tag{4.4}$$

$$L\left(\frac{\|td + \rho(t)\|^2}{t} + 2\varepsilon_1 \left\|d + \frac{\rho(t)}{t}\right\| t^{1/2}\right) \leq \frac{\delta}{6}, \tag{4.5}$$

$$\left\|\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)\right\| \frac{\|\rho(t)\|}{t} \leq \frac{\delta}{6}. \tag{4.6}$$

For each $t > 0$ put

$$\tau(t) = \frac{12\,\mathrm{dist}(F(\sigma(t),\, x(t)),\, Q)}{\delta t}, \tag{4.7}$$

$$\tilde{x}(t) = \tau(t)x_0 + (1 - \tau(t))x(t), \tag{4.8}$$

$$\Phi_1(t) = F(\sigma(t),\, \tilde{x}(t)) - F(\sigma(t),\, x(t)) + \tau(t)\frac{\partial F}{\partial x}(\sigma_0,\, x_0)(x(t) - x_0), \tag{4.9}$$

$$\Phi_2(t) = F(\sigma(t),\, x(t)) - F(\sigma_0,\, x_0) - \frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)(td + \rho(t)) - \frac{\partial F}{\partial x}(\sigma_0,\, x_0)(x(t) - x_0), \quad (4.10)$$

choose an element $p(t) \in Q$ such that

$$\|F(\sigma(t),\, x(t)) - p(t)\| \le 2\,\mathrm{dist}(F(\sigma(t),\, x(t)),\, Q), \quad (4.11)$$

and set

$$q(t) = \tau(t)F(\sigma_0,\, x_0) + (1 - \tau(t))p(t). \quad (4.12)$$

Throughout the rest of the proof we assume that $F(\sigma(t),\, x(t)) \notin Q$ (otherwise estimate (4.3) holds trivially). Then according to (4.2), (4.7) and the definition of ε_2 it holds that

$$0 < \tau(t) = \frac{12\,\mathrm{dist}(F(\sigma(t),\, x(t)),\, Q)}{\delta t} \le \frac{12\varepsilon_2}{\delta} = 1.$$

In particular, by (4.12), $q(t) \in Q$. Furthermore, by (4.2), (4.4) and (4.8) it holds that $\sigma(t) \in B_\varepsilon(\sigma_0)$, $\tilde{x}(t) \in B_\varepsilon(x_0)$.

We next estimate $\|\Phi_1(t)\|$ and $\|\Phi_2(t)\|$ for $t \in (0,\, \bar{t}]$. By (4.1), (4.5), (4.8), (4.9), by the mean value theorem, and by the definition of ε_1 we obtain

$$\begin{aligned}
\|\Phi_1(t)\| &= \Big\| F(\sigma(t),\, x(t) - \tau(t)(x(t) - x_0)) - F(\sigma(t),\, x(t)) \\
&\qquad - \frac{\partial F}{\partial x}(\sigma_0,\, x_0)(-\tau(t)(x(t) - x_0)) \Big\| \\
&\le \sup_{\theta \in [0,\,1]} \left\| \frac{\partial F}{\partial x}(\sigma(t),\, x(t) - \theta\tau(t)(x(t) - x_0)) - \frac{\partial F}{\partial x}(\sigma_0, x_0) \right\| \tau(t)\|x(t) - x_0\| \\
&\le L\left(\|td + \rho(t)\| + \sup_{\theta \in [0,\,1]} \|x(t) - \theta\tau(t)(x(t) - x_0) - x_0\| \right) \tau(t)\|x(t) - x_0\| \\
&\le L\left(\|td + \rho(t)\| + \sup_{\theta \in [0,\,1]} (1 - \theta\tau(t))\|x(t) - x_0\| \right) \tau(t)\|x(t) - x_0\| \\
&\le L\left(\left\| d + \frac{\rho(t)}{t} \right\| t + \|x(t) - x_0\| \right) \tau(t)\|x(t) - x_0\| \\
&\le L\left(\left\| d + \frac{\rho(t)}{t} \right\| t + \varepsilon_1 t^{1/2} \right) \varepsilon_1 \tau(t) t^{1/2} \\
&\le \left(L\varepsilon_1 \left\| d + \frac{\rho(t)}{t} \right\| t^{1/2} + L\varepsilon_1^2 \right) \tau(t) t \\
&\le \left(\frac{\delta}{6} + \frac{\delta}{6} \right) \tau(t) t \\
&= \frac{\delta}{3}\tau(t)t. \qquad (4.13)
\end{aligned}$$

Similarly, by (4.1), (4.5), (4.10), by the mean value theorem, and by the definition of ε_1 we obtain

$$\|\Phi_2(t)\| \le \sup_{\theta \in [0,\,1]} \|F'(\sigma_0 + \theta(td + \rho(t)),\, x_0 + \theta(x(t) - x_0))$$

$$
\begin{aligned}
& -F'(\sigma_0,\, x_0)\|(\|td+\rho(t)\| + \|x(t)-x_0\|) \\
\leq\ & L \sup_{\theta\in[0,\,1]} \theta(\|td+\rho(t)\| + \|x(t)-x_0\|)^2 \\
\leq\ & L(\|td+\rho(t)\|^2 + 2\varepsilon_1\|td+\rho(t)\|t^{1/2} + \varepsilon_1^2 t) \\
=\ & \left(L\left(\frac{\|td+\rho(t)\|^2}{t} + 2\varepsilon_1\left\|d+\frac{\rho(t)}{t}\right\|t^{1/2}\right) + L\varepsilon_1^2\right)t \\
\leq\ & \left(\frac{\delta}{6}+\frac{\delta}{6}\right)t \\
=\ & \frac{\delta}{3}t. \qquad (4.14)
\end{aligned}
$$

We are now in a position to show that for $t \in (0,\, \bar{t}]$

$$
F(\sigma(t),\, \tilde{x}(t)) - q(t) \in \operatorname{cone} B_\delta\left(\frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d\right). \qquad (4.15)
$$

(This will mean that (2.13) is satisfied with $\sigma = \sigma(t)$, $x = \tilde{x}(t)$ and $\bar{y} = -\frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d$.) Indeed, put

$$
y^1(t) = F(\sigma(t),\, x(t)) - p(t), \quad y^2(t) = \frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)\rho(t). \qquad (4.16)
$$

Then according to (4.12), (4.9), (4.10) we have

$$
\begin{aligned}
F(\sigma(t),\, \tilde{x}(t)) - q(t) =\ & F(\sigma(t),\, \tilde{x}(t)) - F(\sigma(t),\, x(t)) + F(\sigma(t),\, x(t)) \\
& +\tau(t)\frac{\partial F}{\partial x}(\sigma_0,\, x_0)(x(t)-x_0) - \tau(t)\frac{\partial F}{\partial x}(\sigma_0,\, x_0)(x(t)-x_0) \\
& -\tau(t)F(\sigma_0,\, x_0) - (1-\tau(t))p(t) \\
=\ & \Phi_1(t) + F(\sigma(t),\, x(t)) - p(t) \\
& -\tau(t)\left(\frac{\partial F}{\partial x}(\sigma_0,\, x_0)(x(t)-x_0) + F(\sigma_0,\, x_0) - p(t)\right) \\
=\ & \Phi_1(t) + y^1(t) \\
& +\tau(t)\left(\Phi_2(t) - F(\sigma(t),\, x(t)) + t\frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d + y^2(t) + p(t)\right) \\
=\ & \tau(t)t\left(\frac{\Phi_1(t)}{\tau(t)t} + \frac{y^1(t)}{\tau(t)t} + \frac{\Phi_2(t)}{t} - \frac{y^1(t)}{t} + \frac{y^2(t)}{t} + \frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d\right) \\
=\ & \tau(t)t\left(\frac{\Phi_1(t)}{\tau(t)t} + (1-\tau(t))\frac{y^1(t)}{\tau(t)t} + \frac{\Phi_2(t)}{t} + \frac{y^2(t)}{t} + \frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d\right),
\end{aligned}
$$

and hence, by (4.6), (4.7), (4.11), (4.13), (4.14), (4.16)

$$
\left\|\frac{F(\sigma(t),\, \tilde{x}(t)) - q(t)}{\tau(t)t} - \frac{\partial F}{\partial\sigma}(\sigma_0,\, x_0)d\right\| \leq \frac{\delta}{3} + \frac{2\delta\operatorname{dist}(F(\sigma(t),\, x(t)),\, Q)}{12\operatorname{dist}(F(\sigma(t),\, x(t))} + \frac{\delta}{3} + \frac{\delta}{6} = \delta,
$$

which proves (4.15).

By the choice of $\varepsilon > 0$, $\delta > 0$ and $c > 0$, the estimate (1.3) holds with $\sigma = \sigma(t)$ and $x = \tilde{x}(t)$. Thus, taking into account (4.7), (4.8), we conclude that $\forall\, t \in (0,\, \bar{t}]$

$$\begin{aligned}
\mathrm{dist}(x(t),\, D(\sigma(t)) &\leq \|x(t) - \tilde{x}(t)\| + \mathrm{dist}(\tilde{x}(t),\, D(\sigma(t)) \\
&\leq \|x(t) - \tilde{x}(t)\| + c\, \mathrm{dist}(F(\sigma(t),\, \tilde{x}(t)),\, Q) \\
&\leq \|x(t) - \tilde{x}(t)\| \\
&\quad + c(\|F(\sigma(t),\, x(t)) - F(\sigma(t),\, \tilde{x}(t))\| + \mathrm{dist}(F(\sigma(t),\, x(t)),\, Q)) \\
&\leq \|x(t) - \tilde{x}(t)\| + c(l\|x(t) - \tilde{x}(t)\| + \mathrm{dist}(F(\sigma(t),\, x(t)),\, Q)) \\
&\leq (1 + c\,l)\|x(t) - x_0\|\tau(t) + c\, \mathrm{dist}(F(\sigma(t),\, x(t)) \\
&\leq \left(c + \frac{12(1 + c\,l)}{\delta}\frac{\|x(t) - x_0\|}{t}\right) \mathrm{dist}(F(\sigma(t),\, x(t)).
\end{aligned}$$

This implies (4.3) with $a = \max\{c,\, 12(1 + c\,l)/\delta\}$. ■

Note that the proof above actually specifies all the constants appearing in the assertion of Lemma 4.1.

The second result is [3, Lemma 4.109]. Our proof is an evident modification of the proof in [1, Lemma 6.2].

Lemma 4.2 *Let Q be closed and convex, and let $x_0 \in D(\sigma_0)$. Let F be Fréchet-differentiable at $(\sigma_0,\, x_0)$ and Fréchet-differentiable with respect to x near $(\sigma_0,\, x_0)$, and let its derivative with respect to x be continuous at $(\sigma_0,\, x_0)$.*

If (1.6) holds at x_0 with respect to a direction $d \in \Sigma$, then there exists $a > 0$ possessing the following property: for any mappings $\rho(\cdot) : \mathbf{R}_+ \to \Sigma$ and $x(\cdot) : \mathbf{R}_+ \to X$ such that $\rho(t) = o(t)$, $x(t) \to x_0$ as $t \to 0$, and the estimate

$$\mathrm{dist}(F(\sigma_0 + td + \rho(t),\, x(t)),\, Q) = o(t) \tag{4.17}$$

holds for $t \geq 0$, and for any $\theta > 0$, the estimate

$$\mathrm{dist}(x(t),\, D(\sigma_0 + (1 + \theta)td + \rho((1 + \theta)t))) \leq a\theta t \tag{4.18}$$

holds for all $t > 0$ small enough.

Proof. By the same argument as in the proof of Lemma 4.1 we can choose $\varepsilon > 0$, $\delta > 0$ and $c > 0$ such that the estimate (1.3) holds for all $(\sigma,\, x) \in B_\varepsilon(\sigma_0) \times B_\varepsilon(\sigma_0)$ satisfying inclusion (2.13) with $\bar{y} = -\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d$. For each $t > 0$ put $\sigma(t) = \sigma_0 + td + \rho(t)$.

For a fixed $\theta > 0$ and for $t \geq 0$ we have

$$F(\sigma((1 + \theta)t),\, x(t)) = F(\sigma(t),\, x(t)) + \theta t \frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d + o(t). \tag{4.19}$$

Select $q(t) \in Q$ such that

$$\|F(\sigma(t),\, x(t)) - q(t)\| = \mathrm{dist}(F(\sigma(t),\, x(t)),\, Q) + o(t).$$

Then for $t > 0$ small enough $\sigma(t) \in B_\varepsilon(\sigma_0)$, $x(t) \in B_\varepsilon(x_0)$, and by (4.17) and (4.19) it holds that

$$\begin{aligned} F(\sigma((1+\theta)t),\, x(t)) - q(t) &= \theta t \frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d + o(t) \\ &\in \operatorname{cone} B_\delta\left(\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d\right), \end{aligned} \tag{4.20}$$

i.e., inclusion (2.13) holds with $\sigma = \sigma((1+\theta)t)$, $x = x(t)$, and with $\bar{y} = -\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d$. Hence by (1.3) and the equality in (4.20) we conclude that

$$\begin{aligned} \operatorname{dist}(x(t),\, D(\sigma((1+\theta)t)) &\leq c \operatorname{dist}(F(\sigma((1+\theta)t),\, Q) \\ &\leq c\|F(\sigma((1+\theta)t),\, x(t)) - q(t)\| \\ &= c\left\|\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d\right\|\theta t + o(t), \end{aligned}$$

and the needed estimate (4.18) holds with any $a > c\|\frac{\partial F}{\partial \sigma}(\sigma_0,\, x_0)d\|$ for all $t > 0$ small enough.
■

Acknowledgment. We would like to thank the anonymous referees for their helpful comments. Proposition 2.2 was entirely motivated by the referee's suggestion to try to remove an extra localization in condition (2.13) of Theorem 2.1 in the original version of the paper. The current version of Theorem 2.1 is of course equivalent to the original one, but its statement and the subsequent presentation became much more elegant. We are specially thankful to the referee for this.

References

[1] A. V. Arutyunov and A. F. Izmailov. Directional stability theorem and directional metric regularity. Math. Oper. Res. 31 (2006), 526–543.

[2] J. F. Bonnans and R. Cominetti. Perturbed optimization in Banach spaces, Part I: A general theory based on a weak directional constraint qualification. SIAM J. Control Optim. 34 (1996), 1151–1171.

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

[4] J. M. Borwein. Stability and regular points of inequality systems. J. Optim. Theory Appl. 48 (1986), 9–52.

[5] J. M. Borwein and D. M. Zhuang. Verifiable necessary and sufficient conditions for openness and regularity of set-valued and single-valued maps. J. Math. Anal. Appl. 134 (1988), 441–459.

[6] A. V. Dmitruk, A. A. Milyutin, and N. P. Osmolovskii. Lyusternik's theorem and the theory of extrema. Russian Math. Surveys. 35 (1980), 11-51.

[7] A. L. Dontchev. The Graves theorem revisited. J. Convex Anal. 3 (1996), 45–54.

[8] A. L. Dontchev, M. Quincampoix, and N. Zlateva. Aubin criterion for metric regularity. J. Convex Anal. 13 (2006), 281–297.

[9] A. L. Dontchev and W. W. Hager. Implicit functions, Lipschitz maps and stability in optimization. Math. Oper. Res. 19 (1994), 753–768.

[10] P. C. Duong and H. Tuy. Stability, surjectivity and local invertibility of nondifferentiable mappings. Acta Math. Vietnam. 3 (1978), 89–105.

[11] B. Gollan. On the marginal function in nonlinear programming. Math. Oper. Res. 9 (1984), 208–221.

[12] L. M. Graves. Some mapping theorems. Duke Math. J. 17 (1950), 111–114.

[13] A. D. Ioffe. Metric regularity and subdifferential calculus. Russian Math. Surveys. 55 (2000), 103–162.

[14] A. D. Ioffe and V. M. Tikhomirov. Theory of Extremal Problems. North–Holland, Amsterdam, The Netherlands, 1979. English-language translation of "Teoriya ekstremal'nikh zadach" (in Russian), Nauka, Noskow, 1974.

[15] L. Lyusternik. Conditional extrema of functions. Math. USSR–Sb. 31 (1934), 390–401.

[16] B. S. Mordukhovich. Variational Analysis and Generalized Differentiation. Springer-Verlag, Berlin, 2005.

[17] B. S. Mordukhovich Complete characterization of opennes, metric regularity and Lipschitzian properties of multifinctions. Trans. Amer. Math. Soc. 340 (1993), 1–36.

[18] B. S. Mordukhovich and B. Wang Restrictive metric regularity and generalized differential calculus in Banach spaces. Inter. J. Maths. Math. Sci. 50 (2004), 2650–2683.

[19] J.-P. Penot. Metric regularity, opennes and Lipschitzian behavior of multifunctions. Nonlinear Anal. 13 (1989), 629–643.

[20] S. M. Robinson. Regularity and stability of convex multivalued functions. Math. Oper. Res. 1 (1976), 130–143.

[21] S. M. Robinson. Stability theorems for systems of inequalities, Part II: Differentiable nonlinear systems. SIAM J. Numer. Anal. 13 (1976), 497–513.

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

[23] C. Ursescu. Multifunctions with closed convex graphs Czech. Math. J. 25 (1975), 438–441.