J. Fixed Point Theory Appl. 17 (2015) 187–208
DOI 10.1007/s11784-015-0240-5
Published online June 3, 2015

Journal of Fixed Point Theory and Applications

On primal regularity estimates for single-valued mappings

Radek Cibulka, Marián Fabian and Alexander D. Ioffe

To Andrzej Granas en tout amitié et estime

Abstract. Based on a primal regularity criterion we provide lower bounds for the regularity modulus of a nonlinear single-valued mapping F from a Banach space X into another Banach space Y. We focus on the case when F is defined on a proper (closed convex) subset of X only rather than on the whole of X. Three possible ways of approximating F around the reference point are considered. First, we use a tangential approximation by set-valued mappings associated with the Bouligand's tangent cone to the graph of F. Then we move on to approximations by positively homogeneous set-valued mappings whose graphs contain the graph of F, for example, by the strict prederivative. Finally, we use an approximation by bunches of continuous linear operators. In the first two cases finding approximating objects is relatively easy while in the third case the approximating object is very convenient to work with. On examples, we illustrate that these approaches are different and neither of them implies the other, unless the spaces in question are finite dimensional.

Mathematics Subject Classification. 49J53, 49J52, 49K40, 90C31.

Keywords. Metric regularity, linear openness, tangential approximation, contingent tangent cone, strict prederivative, measure of noncompactness, Clarke's inverse function theorem, cofinal family, renorming, nonsmooth analysis.

1. Introduction

The three equivalent regularity properties (metric regularity, linear openness and pseudo-Lipschitz property of the inverse) play fundamental role in modern variational analysis concerned basically with set-valued mappings. Necessity to appeal to one of them constantly appears here and there, both in the theory and applications, especially in problems associated with optimization and nonsmooth analysis (see, e.g., [12, 21], see also [24] for a discussion of interrelations of metric regularity and metric fixed point theory).

As far as the regularity theory is concerned, the main question is how to calculate or estimate regularity rates (or moduli) near a certain point of the graph of the mapping. The most universal and often easy to use is the general regularity criterion (see Theorem 2.3 below and the subsequent bibliographic comment) that gives a precise recipe for calculation of the rate of surjection/metric regularity for set-valued mappings between metric spaces. For mappings between Banach spaces (and more generally from a metric space to, say, a length metric space, that is a space in which the distance between any two points coincides with the lower bound of the lengths of curves joining the points) exact expressions for rates of regularity can be obtained in terms of slopes (maximal instantaneous rate of decrease of a function) introduced by De Giorgi, Marino and Tosques in [7]; see [4, 21]. Better known, however, are so-called coderivative estimates based on use of one or another subdifferential constructions (see [19, 20, 21] for the general case and [25] for mappings between Asplund spaces). Popularity of such criteria (although they are often less precise) is based on the fact that they give a direct connection to classical results for smooth maps which the slope-based criterion does not. This makes coderivative criteria convenient in many theoretical arguments. In particular, this is the language that is normally used to state necessary optimality conditions in nonsmooth optimization theory. The problem is that finding coderivatives of concrete mappings often is an unpleasant task associated with heavy calculations and requiring a lot of work.

Here we consider three possible ways of primal estimating rates or regularity. In two of the three cases single-valued (on the domain) mappings are natural objects, so we do not go beyond this class of maps. The common basic idea in all three cases is exactly the same as in the classical calculus: we approximate the mapping by some other object, easier to deal with in one or another sense.

In this paper we consider three types of approximations: (a) tangential approximation by set-valued mappings associated with the tangent cone to the graph of our mapping (with tangency understood in the most general Bouligand sense), (b) approximations by positively homogeneous set-valued mappings whose graphs contain the graph of our mapping and (c) by bunches of linear operators. The advantage of these primal approaches is that in the first two cases finding approximating objects is relatively easy, while in the third case the approximating object is very convenient to work with.

We do not compare here the dual and primal approaches (e.g., which provide better quality of estimates)—this is an interesting and not well-studied question. But it has to be observed that the developments of both started approximately at the same time, in the late 70s and early 80s. Predecessors of all three main results we prove here appeared about that time. But the technical machinery we use, the mentioned general regularity criterion first of all, came into being much later. We shall postpone giving more detailed bibliographic comments till after proofs of the theorems in Section 3.

The plan of the paper is the following. In Section 2 we state several principal facts of the regularity theory of variational analysis we need. All main results are stated and proved in Section 3 along with some consequences and bibliographic comments. In the short final section, Section 4, we discuss the connection between the three main results relating to the three types of approximation we consider: Theorem 3.2 for tangential approximation, Theorem 3.4 for homogeneous approximation and Theorem 3.9 for approximation by sets of linear operators.

Notation and terminology

As a rule, we shall use the same symbol $\|\cdot\|$ to denote norms in different Banach spaces hoping this will cause no confusion and adding whenever necessary a subscript to emphasize which space we are talking about, e.g. $\|\cdot\|_X$. The same stipulation applies to the notation for distance functions, balls etc.: B_X and S_X are, respectively, the closed unit ball and the unit sphere in a Banach space X; $B(x,r)$ is the closed ball centered at $x \in X$ with a radius $r > 0$ and $\mathring{B}(x,r)$ is the corresponding open ball.

The symbol $F : X \rightrightarrows Y$ means that F is a set-valued mapping that may assume the empty value as well. The set $\operatorname{dom} F = \{x : F(x) \neq \emptyset\}$ is the *domain* of F. In this paper we deal with F which are actually single valued on $\operatorname{dom} F$. In this case we say that F is *single valued* (or to avoid confusion, *single-valued on its domain*) and write $F : X \to Y$.

The graph of a mapping F is the set

$$\operatorname{gph} F = \{(x,y) \in X \times Y : y \in F(x)\}$$

and the inverse of F is the mapping

$$Y \ni y \longmapsto \{x \in X : y \in F(x)\} =: F^{-1}(y) \subset X.$$

If both X and Y are Banach spaces, then

$$\mathcal{H} : X \rightrightarrows Y$$

is called *(positively) homogeneous* if $\mathcal{H}(\lambda x) = \lambda \mathcal{H}(x)$ for $\lambda > 0$. The (upper) norm of a homogeneous mapping is

$$\|\mathcal{H}\| = \sup_{\|x\| \le 1} \sup\{\|y\| : y \in \mathcal{H}(x)\}.$$

(The usual convention is that we set

$$\sup \emptyset = -\infty \quad \text{and} \quad \inf \emptyset = \infty.$$

But when we deal with nonnegative quantities, it is more convenient to agree that $\sup \emptyset = 0$.) We say that $\mathcal{H}$ is *bounded* if $\|\mathcal{H}\| < \infty$.

2. Regularity

Given two metric spaces X and Y, a set-valued mapping $F : X \rightrightarrows Y$ is called *open with a linear rate near* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there are $c > 0$ and $\varepsilon > 0$ such that[1]

$$B(y,ct)\cap B(\bar y,\varepsilon)\subset F(B(x,t))\quad \text{if } (x,y)\in B((\bar x,\bar y),\varepsilon)\cap \operatorname{gph}F,\ 0<t<\varepsilon. \tag{2.1}$$

The upper bound of $c>0$ such that (2.1) holds for some $\varepsilon>0$ is called *rate of openness* (or *rate (or modulus) of surjection*) of F near $(\bar x,\bar y)$ and is denoted by $\operatorname{sur}F(\overline{x},\overline{y})$. If F is single valued on its domain, we write $\operatorname{sur}F(\bar x)$ instead of $\operatorname{sur}F(\bar x,F(\bar x))$.

Remark 2.1. As we have mentioned there are two other equivalent characterizations of linear openness known as *metric regularity* and *pseudo-Lipschitz* or *Aubin property*. For instance, $F:X\rightrightarrows Y$ with $(\bar x,\bar y)\in\operatorname{gph}F$ is said to be *metrically regular* near $(\bar x,\bar y)$ if there are $\kappa>0$ and neighborhoods U of $\bar x$ and V of $\bar y$ such that

$$d\big(x,F^{-1}(y)\big)\le \kappa d(y,F(x))\quad \text{for all } (x,y)\in U\times V. \tag{2.2}$$

The lower bound of all $\kappa>0$ such that (2.2) holds for some neighborhoods U and V is called the *rate* or *modulus of metric regularity* of F near $(\bar x,\bar y)$ and is denoted by $\operatorname{reg}F(\overline{x},\overline{y})$. Equivalence of linear openness and metric regularity was first mentioned probably in the 1980 paper by Dmitruk, Milyutin and Osmolowski [10] and formally proved (along with the equivalence with the pseudo-Lipschitz property) by Borwein and Zhuang [5] and Penot [29] in late 80s. It turns out also that the equality

$$\operatorname{reg}F(\bar x,\bar y)\cdot\operatorname{sur}F(\bar x,\bar y)=1$$

always holds (if we set $0\cdot\infty=1$); see [21]. We shall not use metric regularity and the Aubin property in this paper and we simply call F *regular* near $(\bar x,\bar y)$ if it is open with a linear rate near this point.

Theorem 2.2 (Milyutin's perturbation theorem [10]**).** *Let X and Y be metric spaces, let $F:X\rightrightarrows Y$ be a set-valued mapping with closed graph, let $\overline{x}\in\operatorname{dom}F$, $\overline{y}\in F(\overline{x})$, and let $G:X\to Y$ be defined and Lipschitz in a neighborhood of $\overline{x}$. Then*

$$\operatorname{sur}(F+G)(\overline{x},\overline{y}+G(\overline{x}))\ge\operatorname{sur}F(\overline{x},\overline{y})-\operatorname{lip}G(\overline{x}).$$

Here $\operatorname{lip}G(\overline{x})$ is the Lipschitz constant of G at $\overline{x}$, that is, the lower bound of Lipschitz constants of G on neighborhoods of $\overline{x}$.

The following regularity criterion plays a central role in our proofs.

Theorem 2.3 (General criterion for single-valued maps). *Let X be a complete metric space, let Y be a metric space, and let $F:X\to Y$ be a mapping with closed graph which is continuous on its domain. Let finally $\overline{x}\in\operatorname{dom}F$. Then F is open near $\overline{x}$ with $\operatorname{sur}F(\overline{x})\ge c$ if and only if for any $c'<c$ there is a neighborhood U of $\overline{x}$ such that for any $x\in U\cap\operatorname{dom}F$ and any $y\ne F(x)$ there is an $x'\in\operatorname{dom}F$ such that*

$$d(F(x'),y)<d(F(x),y)-c'd(x',x). \tag{2.3}$$

[1]There are several equivalent definitions of openness near a point in the literature; see, e.g., [22].

This is a special case of a more general criterion applied to arbitrary set-valued mappings. Both Theorem 2.3 and its set-valued extension first appeared in [15] (Corollary 1 and Remark 2c) as by-products of the main result but did not attract much attention at that time. The results were rediscovered in [21] and recognized as an extremely powerful instrument in the regularity theory of variational analysis. We can refer the reader to [22] for various versions of the criterion, local and nonlocal, and demonstrations of its work and efficiency. We shall need the following consequence of the theorem (that goes back to [10, 31, 32] and, in fact, to the proof of the Banach open mapping theorem).

Theorem 2.4 (Density theorem [10, 22]**).** *Let X and Y be complete metric spaces, and let F be a closed-graph single-valued mapping from X into Y which is continuous on its domain. Let $\overline{x} \in \operatorname{dom} F$ and $\overline{y} = F(\overline{x})$. Suppose that there are $c > 0$ and $\varepsilon > 0$ such that $F(B(x,t))$ is dense in*

$$B(F(x), ct) \cap B(\overline{y}, \varepsilon)$$

for all $x \in \operatorname{dom} F$ close to $\overline{x}$ and all $t \in (0, \varepsilon)$. Then $\operatorname{sur} F(\overline{x}) \geq c$.

For Banach spaces (and more generally for metric spaces with Y, the range space, being the length space) it is possible to prove an infinitesimal version of the general criterion which we shall also formulate here for single-valued mappings (see [4, 21]).

Given a function f on a metric space X which is finite at $x \in X$. The quantity

$$|\nabla f|(x) = \limsup_{x \neq u \to x} \frac{(f(x) - f(u))^+}{d(x,u)}$$

is called the *slope* of f at x. Here as usual, $\alpha^+ = \max\{\alpha, 0\}$.

Let a mapping $F : X \to Y$ be given. Throughout the paper we shall use the notation

$$\varphi_y(x) = \begin{cases} \|y - F(x)\| & \text{if } x \in \operatorname{dom} F; \\ \infty & \text{otherwise.} \end{cases} \tag{2.4}$$

We also fix some $\overline{x} \in \operatorname{dom} F$ and set $\overline{y} = F(\overline{x})$.

Theorem 2.5. *Let X and Y be Banach spaces, let $F : X \to Y$ be a mapping with closed graph which is continuous on its domain. Let finally $\overline{x} \in \operatorname{dom} F$ and let the function φ_y be defined by (2.4). Then F is open near $\overline{x}$ with $\operatorname{sur} F(\overline{x}) \geq c$ if and only if for any $c' < c$ there is a neighborhood U of $\overline{x}$ such that for any $x \in U \cap \operatorname{dom} F$ and any $y \neq F(x)$,*

$$|\nabla \varphi_y|(x) \geq c'.$$

3. Main results

Here and below we adopt the following basic hypothesis: *X and Y are Banach spaces and F is a single-valued mapping from X into Y which is continuous* *on its domain and such that for any* $w \in \mathrm{cl}(\mathrm{dom}\, F)\setminus \mathrm{dom}\, F$,

$$\mathrm{dom}\, F \ni x \to w \Longrightarrow \|F(x)\| \to \infty. \tag{3.1}$$

Under this assumption the function φ_y in (2.4) is lower semicontinuous.

3.1. Tangential approximations

Given a set $S \subset X$ and an $\overline{x} \in S$. The *contingent tangent cone* $T(S, \overline{x})$ is the collection of $h \in X$ with the following property: there are sequences of $t_k \downarrow 0$ and $h_k \to h$ such that $\overline{x} + t_k h_k \in S$ for all k.

Let now f be a function on X finite at $\overline{x}$. The function

$$X \ni h \longmapsto f^-(\overline{x}; h) := \liminf_{(t,h')\to(0+,h)} t^{-1}\big(f(\overline{x} + th') - f(\overline{x})\big)$$

is called the *Dini–Hadamard lower directional derivative* of f at $\overline{x}$. This function is either lower semicontinuous and equal to zero at the origin or identically equal to $-\infty$. The latter of course cannot happen if f is Lipschitz near $\overline{x}$.

The connection between the two concepts is very simple: $h \in T(S, \overline{x})$ if and only if $d^-(\cdot, S)(\overline{x}; h) = 0$, and $\alpha = f^-(\overline{x}; h)$ if and only if (h, α) is in the tangent cone to $\mathrm{gph}\, f$ at $(\overline{x}, f(\overline{x}))$.

The following simple proposition establishes connection between the slope of f and its lower directional derivative.

Proposition 3.1. *For any function f and any x at which f is finite,*

$$|\nabla f|(x) \geq -\inf_{\|h\|=1} f^-(x; h).$$

Proof. Take an h with $\|h\| = 1$. We have

$$\begin{aligned}
|\nabla f|(x) &= \lim_{t\downarrow 0} \sup_{\|u\|=1} \frac{(f(x) - f(x+tu))^+}{t} \\
&\geq \limsup_{(t,u)\to(0+,h)} \frac{f(x) - f(x+tu)}{t} \\
&= -f^-(x; h)
\end{aligned}$$

as claimed. □

If $F : X \to Y$, then the *contingent derivative* of F at $\overline{x} \in \mathrm{dom}\, F$ is the set-valued mapping

$$X \ni h \longmapsto DF(\overline{x}; h) := \big\{v \in Y : (h, v) \in T(\mathrm{gph}\, F, (\overline{x}, F(\overline{x})))\big\}.$$

We are ready to state the main result of this subsection.

Theorem 3.2. *Under the basic hypothesis,*

$$\mathrm{sur}\, F(\overline{x}) \geq \alpha \tag{3.2}$$

if for any $\alpha' < \alpha$ there are neighborhoods $U \subset X$ of $\overline{x}$ and $V \subset Y$ of $\overline{y}$ such that for any $x \in U \cap \mathrm{dom}\, F$ and any $V \ni y \neq F(x)$ there is an $h \in S_X$ such that $\varphi_y^-(x; h) \leq -\alpha'$, where the function φ_y is defined by (2.4).

In particular, suppose that there exist $c > 0$ *and* $\lambda \in [0, 1)$ *such that for any* $c' > c$ *and* $\lambda' > \lambda$ *there is a neighborhood* U *of* $\overline{x}$ *such that for any* $x \in U \cap \operatorname{dom} F$ *and any* $z \in S_Y$ *there is an* $h \in X$ *with* $\|h\| \leq c'$ *and* $d(z, DF(x; h)) < \lambda'$*. Then*

$$\operatorname{sur} F(\overline{x}) \geq \frac{1 - \lambda}{c}. \tag{3.3}$$

Proof. The first statement is immediate from Proposition 3.1 and Theorem 2.5.

To prove the second, take $c' > c$ and $\lambda' \in (\lambda, 1)$ and find a corresponding neighborhood U of $\overline{x}$. Take an $x \in U \cap \operatorname{dom} F$ and $y \neq F(x)$ and set

$$z := \|y - F(x)\|^{-1}(y - F(x)).$$

By the assumption, there is an $h \in X$ with $\|h\| \leq c'$ such that $\|z - v\| < \lambda'$ for some $v \in DF(x; h)$. The latter means that there are $t_k \downarrow 0$ and $h_k \to h$, $k = 1, 2, \ldots$, such that $v_k = t_k^{-1}(F(x + t_k h_k) - F(x)) \to v$. We have

$$\begin{aligned} \|y - F(x + t_k h_k)\| &\leq \|y - F(x) - t_k z\| + t_k \|z - v_k\| \\ &= \|y - F(x)\| - t_k + t_k \|z - v_k\|. \end{aligned}$$

Therefore,

$$\varphi_y^-(x; h) \leq \liminf_{k \to \infty} t_k^{-1}\big(\|y - F(x + t_k h_k)\| - \|y - F(x)\|\big) \leq -(1 - \lambda').$$

If $h = 0$, it follows that $\varphi_y^-(x; 0) = -\infty$. If $h \neq 0$, we get

$$\varphi_y^-\left(x; \frac{h}{\|h\|}\right) \leq -\frac{(1 - \lambda')}{c'}$$

and it remains to apply the first statement taking into account that c' and λ' can be arbitrarily close to c and λ. □

The theorem can easily be extended to set-valued mappings with practically the same proof (based on the set-valued version of Theorem 2.5). The (set-valued version of the) second statement of the theorem is a long known result. Its "qualitative" part, namely linear openness of the mapping near the nominal point (with a somewhat less precise estimate) was proved by Aubin in [2] in 1981 (see also [3]). The estimate (3.3) was obtained six years later in [20]. Closely connected with Aubin's theorem is the result of Dontchev, Quincampoix and Zlateva [11]. As to the first part of the theorem, it seems to appear for the first time.

It has to be observed that the criterion provided by the first statement is strictly stronger (unless both spaces are finite dimensional). Informally, this is easy to understand: the quality of approximation provided by the contingent derivative for a map into an infinite-dimensional space may be much lower than for a real-valued function. The following example illustrates the phenomenon.

Example 3.3. Let $X = Y$ be a separable Hilbert space, and let $(e_1, e_2, \dots)$ be an orthonormal basis in X. Consider the following mapping from $[0,1]$ into X:

$$\eta(t) = \begin{cases} 0 & \text{if } t \in \{0,1\}, \\ 2^{-(n+2)} e_n & \text{if } t = 2^{-n},\ n = 1,2,\dots, \end{cases}$$

and $\eta(\cdot)$ is linear on every segment $[2^{-(n+1)}, 2^{-n}]$, $n = 0,1,\dots$. Then

$$2\left\|\eta\left(\frac{1}{2}\right) - \eta(1)\right\| \le \frac{1}{4}$$

and, for any $n \in \mathbb{N}$, we have

$$\begin{aligned} \frac{\|\eta(2^{-(n+1)}) - \eta(2^{-n})\|}{|2^{-(n+1)} - 2^{-n}|} &= 2^{n+1}\|2^{-(n+3)}e_{n+1} - 2^{-(n+2)}e_n\| \\ &= \frac{\sqrt{5}}{4} < \frac{3}{4}. \end{aligned} \tag{3.4}$$

Define a mapping from the unit ball of ℓ_2 into ℓ_2 by

$$F(x) = x - \eta(\|x\|).$$

Using (3.4) we get that $x \mapsto \eta(\|x\|)$ is $(3/4)$-Lipschitz, hence by Milyutin's perturbation theorem F is open near the origin with the rate of surjection at least $1/4$.

Let us look what we get applying both statements of Theorem 3.2 for the mapping. If $\|h\| = 1$ and $t \in [2^{-(n+1)}, 2^{-n}]$, then

$$F(th) = th - (2^{n+1}t - 1)2^{-(n+2)}e_n - (2 - 2^{n+1}t)2^{-(n+3)}e_{n+1};$$

that is,

$$t^{-1}F(th) = h - 2^{-1}e_n + 2^{-2}e_{n+1} + 2^{-(n+2)}t^{-1}(e_n - e_{n+1}).$$

Thus $t^{-1}F(th)$ does not converge when t goes to zero. Hence the contingent cone to the graph of F at zero consists of a single point $(0,0)$ and the second statement in Theorem 3.2 gives $\operatorname{sur} F(0) \ge 0$—a trivial conclusion.

Now take an x with $\|x\| < 1$ and a $y \ne F(x)$. For any $t > 0$ and $h \in X$ such that $\|x + th\| \le 1$, we have

$$\begin{aligned} \|F(x+th) - y\| &= \|x + th - \eta(\|x+th\|) - y\| \\ &\le \|x + th - \eta(\|x\|) - y\| + \|\eta(\|x+th\|) - \eta(\|x\|)\| \\ &\le \|F(x) + th - y\| + \frac{3}{4}t\|h\|. \end{aligned}$$

Taking $h = (y - F(x))/\|y - F(x)\|$, we get

$$\begin{aligned} \varphi_y^-(x;h) &\le \lim_{t\downarrow 0} t^{-1}\left(\left(1 - \frac{t}{\|F(x) - y\|}\right)\|F(x) - y\| - \|F(x) - y\|\right) + \frac{3}{4} \\ &= -\frac{1}{4} \end{aligned}$$

which gives $\operatorname{sur} F(x) \ge 1/4$ for all x with $\|x\| < 1$.

3.2. Homogeneous approximation

In this subsection we deduce the openness of a mapping F around the reference point from the properties of a certain positively homogeneous set-valued mapping. Given a subset S of X, the cone generated by S is denoted by $\operatorname{cone} S$; that is, $\operatorname{cone} S := [0, \infty)S$.

Theorem 3.4. *Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be Banach spaces, and let $F : X \to Y$ be continuous on its domain which is assumed to be closed and convex. Assume that for a given $\bar{x} \in \operatorname{dom} F$ there are positive constants ρ, β and r along with a positively homogeneous set-valued mapping $\mathcal{H} : X \rightrightarrows Y$ such that*

$$F(x') - F(x) \in \mathcal{H}(x' - x) + \beta\|x' - x\|B_Y \quad \forall x, x' \in \operatorname{dom} F \cap B(\bar{x}, r), \tag{3.5}$$

and for all $x \in \operatorname{dom} F$ sufficiently close to $\overline{x}$ and all $y^* \in S_{Y^*}$ we have

$$\sup_{h \in \operatorname{cone}(\operatorname{dom} F - x) \cap S_X} \inf_{w \in \mathcal{H}(h)} \langle y^*, w\rangle \geq \beta + \rho. \tag{3.6}$$

Assume finally that one of the following two conditions holds:

(a) *the norm in Y is Gâteaux smooth and $\mathcal{H}$ has relatively norm compact values;*
(b) *the norm in Y is Fréchet smooth and $\mathcal{H}$ has bounded values.*

Then $\operatorname{sur} F(\overline{x}) \geq \rho$.

Proof. So assume that (a) holds. Take an $\varepsilon \in (0, \rho/2)$. Find $\gamma \in (0, r/2)$ so that (3.6) holds for $x \in B(\overline{x}, \gamma) \cap \operatorname{dom} F$. Take such an x and any $y \in Y$, different from $F(x)$. Let y^* denote the derivative of $\|\cdot\|$ at $y - F(x)$. Then

$$\lim_{0 \neq t \to 0} t^{-1}\big(\|y - F(x) + tw\| - \|y - F(x)\|\big) - \langle y^*, w\rangle = 0 \quad \text{for every } w \in Y. \tag{3.7}$$

By (3.6), there is an $h \in \operatorname{cone}(\operatorname{dom} F - x) \cap S_X$ such that

$$\langle y^*, w\rangle > \beta + \rho - \varepsilon \quad \forall w \in \mathcal{H}(h). \tag{3.8}$$

Fix $\delta > 0$ such that $h \in \delta(\operatorname{dom} F - x)$. Since the set $-\mathcal{H}(h)$ is relatively compact and the limit in (3.7) is uniform with respect to w from any fixed compact set, we get that there is $t \in (0, \min\{1/\delta, \gamma\})$ such that

$$\|y - F(x) - tw\| - \|y - F(x)\| + \langle y^*, tw\rangle < t\varepsilon \quad \forall w \in \mathcal{H}(h).$$

This and (3.8) imply that

$$\begin{aligned} \|y - F(x) - tw\| &< \|y - F(x)\| - \langle y^*, tw\rangle + \varepsilon t \\ &\leq \|y - F(x)\| - t(\beta + \rho - 2\varepsilon) \quad \forall w \in \mathcal{H}(h). \end{aligned} \tag{3.9}$$

Let $x' := x + th$. Noting that $t \in (0, 1/\delta)$, we have $x' \in x + t\delta(\operatorname{dom} F - x) \subset \operatorname{dom} F$ by convexity of $\operatorname{dom} F$. As $2t < 2\gamma < r$, we have $\|x' - x\| = \|th\| = t < r/2$. Thus $x' \in \operatorname{dom} F \cap B(\bar{x}, r)$. Since $\mathcal{H}$ is positively homogeneous, we have $\mathcal{H}(x' - x) = \mathcal{H}(th) = t\mathcal{H}(h)$. Thus by (3.5) there is a $w \in \mathcal{H}(h)$ such that

$$\|F(x') - F(x) - tw\| \leq \beta t. \tag{3.10}$$

Now, we are ready for the following chain of estimates:

$$\begin{aligned}\|y - F(x')\| &\le \big\|F(x) - F(x') + tw\big\| + \big\|y - F(x) - tw\big\| \\ &< \beta t + \|y - F(x)\| - (\beta + \rho - 2\varepsilon)t \qquad \text{(by (3.10) and (3.9))} \\ &= \|y - F(x)\| - (\rho - 2\varepsilon)t \\ &= \|y - F(x)\| - (\rho - 2\varepsilon)\|x' - x\|.\end{aligned}$$

It remains to apply the criterion of Theorem 2.3.

If (b) holds, the proof is the same. We only have to take into account that under (b) the limit in (3.7) is uniform for w's from any bounded subset of Y. □

There is a canonical way of constructing a positively homogeneous mapping $\mathcal{H}$ associated with given mapping and a point in its domain. It is associated with the concept of "strict prederivative" introduced in [19]: a positively homogeneous set-valued mapping $\mathcal{H} : X \rightrightarrows Y$ is a *strict prederivative* of F at $\overline{x} \in \operatorname{dom} F$ if

$$F(x + h) - F(x) \in \mathcal{H}(h) + r(x, h)B_Y,$$

where $\|h\|^{-1}r(x, h) \to 0$ as $x \to \overline{x}$, $h \to 0$ (and of course both x and $x + h$ belong to $\operatorname{dom} F$).

To construct a strict prederivative that ensures a reasonable "outer" approximation for F, take an $\varepsilon > 0$ and set

$$\begin{aligned}\mathcal{H}_\varepsilon(h) := \big\{&\lambda^{-1}(F(x + \lambda h) - F(x)) : \\ &x,\, x + \lambda h \in \operatorname{dom} F \cap B(\bar{x}, \varepsilon),\, \lambda > 0\big\}, \quad h \in X.\end{aligned}$$

Then $0 \in \mathcal{H}_\varepsilon(0)$ and for $t > 0$ we have

$$\begin{aligned}\mathcal{H}_\varepsilon(th) = t\big\{&(t\lambda)^{-1}(F(x + t\lambda h) - F(x)) : \\ &x,\, x + t\lambda h \in \operatorname{dom} F \cap B(\bar{x}, \varepsilon),\, \lambda > 0\big\},\end{aligned}$$

that is, $\mathcal{H}_\varepsilon(th) = t\mathcal{H}_\varepsilon(h)$. Thus $\mathcal{H}_\varepsilon$ is positively homogeneous and it is an easy matter to see that (3.5) holds with $\beta = 0$.

We get an immediate corollary of the theorem above.

Corollary 3.5. *Assume that Y is a Gâteaux smooth Banach space. Let $F : X \to Y$ satisfy the basic hypothesis and have a closed convex domain. If $\mathcal{H}$ is a strict prederivative of F at $\overline{x}$ with relatively compact values and* (3.6) *holds (with $\beta := 0$), then* $\operatorname{sur} F(\overline{x}) \ge \rho$.

Part (a) of Theorem 3.4 can be equivalently reformulated in somewhat more general terms. Given a set S in a Banach space, the *measure of noncompactness* of S is the lower bound of $r > 0$ such that S can be covered by finitely many open balls of radius r; see [1]. We denote the measure of noncompactness of S by $\chi(S)$.

Theorem 3.6. *Let $F : X \to Y$ be as in Theorem* 3.4. *Assume that Y has Gâteaux smooth norm and $\mathcal{H} : X \rightrightarrows Y$ verifies* (3.5) *and* (3.6) *for some* *positive constants* ρ, β *and* r. *If there is a* $\gamma \geq 0$ *such that* $\chi(\mathcal{H}(x)) \leq \gamma$ *for all* $x \in S_X \cap \operatorname{dom} \mathcal{H}$, *then* $\operatorname{sur} F(\overline{x}) \geq (\rho - \gamma)^+$.

Proof. If $\rho \leq \gamma$, the statement is trivial, so we assume that $\rho > \gamma$. Take a $\delta \in (\gamma, \rho)$. For any $x \in S_X \cap \operatorname{dom} \mathcal{H}$ choose a finite set $\mathcal{H}_0(x) \subset \mathcal{H}(x)$ such that $\mathcal{H}(x) \subset \mathcal{H}_0(x) + \delta B_Y$. Set further for any $x \in \operatorname{dom} \mathcal{H}$,

$$\mathcal{H}_1(x) = \begin{cases} 0 & \text{if } x = 0; \\ \|x\| \mathcal{H}_0(x/\|x\|) & \text{otherwise.} \end{cases}$$

Then $\mathcal{H}_1$ is positively homogeneous. Both (3.5) and (3.6) hold with $\mathcal{H}$ replaced by $\mathcal{H}_1$, β by $\beta + \delta$ and ρ by $\rho - \delta$. Apply Theorem 3.4 to get that $\operatorname{sur} F(\overline{x}) \geq \rho - \delta$. Letting $\delta \downarrow \gamma$ we conclude the proof. □

Corollary 3.5 and, all the more, Theorems 3.4 and 3.6 substantially improve the earlier results in the same vein proved in [19]. We note in this connection that condition (3.6) is equivalent to

$$\liminf_{\operatorname{dom} F \ni x \to \bar{x}} C\big(\mathcal{H}, \operatorname{cone}(\operatorname{dom} F - x)\big) \geq \beta + \rho,$$

where, for a given positively homogeneous $\mathcal{H} : X \rightrightarrows Y$ and a nonempty cone $K \subset X$,

$$C(\mathcal{H}, K) := - \sup_{y^* \in S_{Y^*}} \inf_{h \in K \cap B_X} \sup_{w \in \mathcal{H}(h)} \langle y^*, w \rangle,$$

where $C(\mathcal{H}, K)$ is the Banach constant of $\mathcal{H}$ on K introduced in a somewhat different form in [19]. Páles [27] was the first to use measures of noncompactness in a similar context but in connection with prederivatives defined by bunches of linear operators that will be considered in the next subsection.

3.3. Approximations by sets of linear operators

If $\mathcal{T}$ is a collection of linear operators from X to Y, then the set-valued mapping $X \ni x \mapsto \mathcal{H}(x) := \{Tx : T \in \mathcal{T}\}$ is of course positively homogeneous. We shall consider positively homogeneous mappings defined by sets of bounded linear operators, that is, elements of the space $\mathcal{L}(X, Y)$ endowed with the norm $\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\}$. It is an easy matter to see that in this case $\mathcal{H}$ inherits some properties of $\mathcal{T}$: if $\mathcal{T}$ is (relatively) norm compact in $\mathcal{L}(X, Y)$, then so are the values of $\mathcal{H}$, if $\mathcal{T}$ is bounded, then the values of $\mathcal{H}$ are also bounded.

This observation offers an easy way to specify Theorems 3.4 and 3.6 for $\mathcal{H}$ defined by a set of linear operators. We, however, shall be interested in a somewhat different question, namely which properties of individual operators of $\mathcal{T}$ allow to get conclusions similar to those of the theorems.

To this end we prove the following result.

Proposition 3.7. *Let* $(X, \|\cdot\|)$ *and* $(Y, \|\cdot\|)$ *be Banach spaces, let* $0 \in K \subset X$ *be a convex closed set, let* $y^* \in S_{Y^*}$, *let* $\alpha > 0$ *and let* $\mathcal{T} \subset \mathcal{L}(X, Y)$ *be a convex set such that*

$$T(B_X \cap K) \supset \alpha B_Y \quad \textit{for every } T \in \mathcal{T}. \tag{3.11}$$

(a) *Suppose that the set* $C := \{T^*y^* : T \in \mathcal{T}\}$ *is weak** *closed and* $\mathcal{T}$ *is bounded in* $\mathcal{L}(X,Y)$. *Then, given an* $\alpha' \in (0,\alpha)$, *there exists a nonzero* $h \in K$ *such that*

$$\langle y^*, Th\rangle \geq \alpha'\|h\| \quad \text{for every } T \in \mathcal{T}.$$

(b) *If* X *is reflexive, then there exists a nonzero* $h \in K$ *such that*

$$\langle y^*, Th\rangle \geq \alpha\|h\| \quad \text{for every } T \in \mathcal{T}.$$

Proof. (a) Set

$$D := \Big\{x^* \in X^* : \sup_{x \in B_X \cap K} \langle x^*, x\rangle \leq \alpha'\Big\}.$$

This is a weak* closed convex set, disjoint from C. Indeed, if T^*y^* were in D for some $T \in \mathcal{T}$, we would have from (3.11) that

$$\alpha' \geq \sup_{x \in B_X \cap K} \langle T^*y^*, x\rangle = \sup_{x \in B_X \cap K} \langle y^*, Tx\rangle > \sup_{y \in \alpha' B_Y} \langle y^*, y\rangle = \alpha'\|y^*\| = \alpha',$$

a contradiction.

Since $\mathcal{T}$ is bounded, C is weak* bounded, and so weak* compact. Thus, we can separate C and D by an element of X (see, e.g., [13, Theorem V.2.10]); that is, there is an $h \in X$ strongly separating C and D. Multiplying h by a constant (if there is a need), we can guarantee that

$$0 \leq \sup_{x^* \in D} \langle x^*, h\rangle \leq \alpha' < \inf_{x^* \in C} \langle x^*, h\rangle. \tag{3.12}$$

We claim that $h \in B_X \cap K$. If not, then [13, Corollary V.2.12] yields an $x^* \in X^*$ such that $\langle x^*, h\rangle > \sup_{x \in B_X \cap K} \langle x^*, x\rangle$. Multiplying x^* by a positive constant, if necessary, we may assume that

$$\langle x^*, h\rangle > \alpha' \geq \sup_{x \in B_X \cap K} \langle x^*, x\rangle.$$

Then $x^* \in D$ and (3.12) implies that $\langle x^*, h\rangle \leq \alpha'$; a contradiction. Therefore, $h \in B_X \cap K$.

Finally, as $D \supset \alpha' B_{X^*}$, (3.12) implies that

$$\inf_{x^* \in C} \langle x^*, h\rangle \geq \sup_{x^* \in D} \langle x^*, h\rangle \geq \sup_{x^* \in \alpha' B_{X^*}} \langle x^*, h\rangle = \alpha'\|h\|.$$

(b) If X is reflexive, we set

$$D := \Big\{x^* \in X^* : \sup_{x \in B_X \cap K} \langle x^*, x\rangle < \alpha\Big\}$$

and apply [13, Theorem V.2.8] to justify the existence of an h separating C and D. The subsequent arguments are essentially the same as above. □

If $(Y, \|\cdot\|)$ is reflexive, then Y^* has an equivalent locally uniformly rotund norm, say $|\cdot|$, see [8, Theorem VII.1.14]; then given an $\varepsilon > 0$, the new norm $Y^* \ni y^* \mapsto (\|y^*\|^2 + \varepsilon|y^*|^2)^{1/2}$ is still equivalent and locally uniformly rotund; finally Shmulyan's test guarantees that the corresponding predual norm on Y will be Fréchet smooth, and not far from the original norm. Therefore, combining Proposition 3.7 with the second statement of Theorem 3.4 we get the following result.

Theorem 3.8. *Assume that X is a reflexive space and the mapping $F : X \to Y$ has closed convex domain and is continuous on its domain. Let $\overline{x} \in \operatorname{dom} F$, and let $\mathcal{T} \subset \mathcal{L}(X, Y)$ be a bounded convex set of linear operators such that for some $\rho > 0$ and $\beta > 0$:*

(a) *for any $x, x' \in \operatorname{dom} F$ in a neighborhood of $\overline{x}$ there is a $T \in \mathcal{T}$ such that*

$$\|F(x) - F(x') - T(x - x')\| \leq \beta\|x - x'\|; \tag{3.13}$$

(b) *there is an $\varepsilon > 0$ such that for any $T \in \mathcal{T}$*

$$\varepsilon(\rho + \beta)B_Y \subset T((\varepsilon B_X) \cap (\operatorname{dom} F - \overline{x})). \tag{3.14}$$

Then $\operatorname{sur} F(\overline{x}) \geq \rho$.

Proof. Without any loss of generality, assume that $\overline{x} = 0$ and $F(\overline{x}) = 0$. Since X is reflexive, assumption (b) implies that so is Y. Indeed, fix any $T \in \mathcal{T}$. As T is surjective, [14, Corollary 2.26(iii)] says that Y is isomorphic to $X/T^{-1}(0)$. The continuity of T implies that $T^{-1}(0)$ is the closed subspace of X. Thus $X/T^{-1}(0)$ is reflexive by [14, Exercise 3.114]. Then Y is reflexive by [14, Exercise 3.112]. By the reasoning before the theorem, we can also suppose that the norm on Y is Fréchet smooth.

Let $\rho' \in (0, \rho)$ be arbitrary. Pick $\gamma \in (0, 1)$ such that

$$\beta + \rho' < (1 - \gamma)(\beta + \rho).$$

As $\mathcal{T}$ is bounded, there is a constant $r > 0$ such that for each $x \in rB_X$ and each $T \in \mathcal{T}$ we have

$$(1 - \gamma)B_X - \varepsilon^{-1}x \subset B_X \quad \text{and} \quad \|Tx\| \leq \varepsilon\big((1 - \gamma)(\beta + \rho) - \beta - \rho'\big).$$

Shrink r, if necessary, so that for each $x, x' \in (rB_X) \cap \operatorname{dom} F$ there is an operator $T \in \mathcal{T}$ such that (3.13) holds; that is, (3.5) holds for

$$\mathcal{H}(x) := \{Tx : T \in \mathcal{T}\}, \quad x \in X.$$

Fix any $x \in (rB_X) \cap \operatorname{dom} F$ and any $y^* \in S_{Y^*}$. For a given $T \in \mathcal{T}$, the convexity of $\operatorname{dom} F$ and (3.14) imply that

$$\begin{aligned} T\big((\varepsilon B_X) \cap (\operatorname{dom} F - x)\big) &\supset T\big((\varepsilon(1 - \gamma)B_X) \cap \operatorname{dom} F - x\big) \\ &\supset T\big((1 - \gamma)\big[(\varepsilon B_X) \cap \operatorname{dom} F\big]\big) - Tx \\ &\supset (1 - \gamma)\varepsilon(\beta + \rho)B_Y - Tx \supset \varepsilon(\beta + \rho')B_Y. \end{aligned}$$

Proposition 3.7(b) implies that there is a nonzero $h \in \varepsilon^{-1}(\operatorname{dom} F - x)$ such that

$$\inf_{T \in \mathcal{T}} \langle y^*, Th\rangle \geq (\beta + \rho')\|h\|;$$

that is, (3.6) holds for $\mathcal{H}$ with ρ replaced by ρ'.

Apply Theorem 3.4(b) to get that $\operatorname{sur} F(0) \geq \rho'$. Taking $\rho' \uparrow \rho$ we finish the proof. □

In the same way we could prove an operator version of the first part of Theorem 3.4. But with a compact set of linear operators a stronger result is available. The following theorem is the main result of this subsection.

Theorem 3.9. *Let X and Y be Banach spaces, and let $F : X \to Y$ be a continuous mapping with closed convex domain. Assume that for a given $\overline{x} \in \operatorname{dom} F$ there is a convex subset $\mathcal{T}$ of $\mathcal{L}(X,Y)$ which is relatively compact in $\mathcal{L}(X,Y)$ and that conditions* (a) *and* (b) *in Theorem* 3.8 *hold for some $\rho > 0$ and $\beta > 0$. Then $\operatorname{sur} F(\overline{x}) \geq \rho$.*

We need the following two lemmas to furnish the proof. The first lemma is a sort of a weak lifting result for the linear openness/metric regularity property (see [23]). Let $L \in \mathcal{S}(X)$ and $M \in \mathcal{S}(Y)$, and let $F : X \rightrightarrows Y$. We denote by $F_{L\times M}$ the set-valued mapping from L into M whose graph coincides with $\operatorname{gph} F \cap (L \times M)$ and by $\mathcal{S}(X)$ and $\mathcal{S}(Y)$ the collections of separable subspaces of X and Y, respectively.

Lemma 3.10. *Let $F : X \rightrightarrows Y$ be a mapping with $(\overline{x}, \overline{y}) \in \operatorname{gph} F$. Suppose that there is a family $\mathcal{E}$ of pairs (L, M), where $L \in \mathcal{S}(X)$ and $M \in \mathcal{S}(Y)$ are closed subspaces containing, respectively, $\overline{x}$ and $\overline{y}$. We assume that $\mathcal{E}$ is cofinal*[2] *in $\mathcal{S}(X) \times \mathcal{S}(Y)$ and that the set-valued mapping $F_{L\times M}$ is linearly open near $(\bar{x}, \bar{y})$ with $\operatorname{sur} F_{L\times M}(\bar{x}, \bar{y}) \geq r$ for any pair $(L, M) \in \mathcal{E}$. Then $\operatorname{sur} F(\bar{x}, \bar{y}) \geq r$.*

Proof. We have to show that $\operatorname{sur} F(\bar{x}, \bar{y}) \geq r - \varepsilon$ for any $\varepsilon > 0$. Assuming the contrary, we can find an $\varepsilon \in (0, r)$ and sequences (x_n) in $\operatorname{dom} F$, (v_n) in Y and (y_n) in Y converging, respectively, to $\overline{x}$ and $\overline{y}$ and such that $v_n \in F(x_n)$ and $y_n \notin F(B(x_n, t_n))$, where $t_n = d(v_n, y_n)/(r - \varepsilon)$ for every $n \in \mathbb{N}$.

Now let $L_0 \in \mathcal{S}(X)$ and $M_0 \in \mathcal{S}(Y)$ be such that $x_n \in L_0$ and $v_n, y_n \in M_0$ for every $n \in \mathbb{N}$. Then we have to conclude that $y_n \notin F_{L\times M}(B_L(x_n, t_n))$ for any subspaces $L \subset X$ and $M \subset Y$ containing, respectively, L_0 and M_0, and therefore $\operatorname{sur} F_{L\times M}(\bar{x}, \bar{y}) \leq r - \varepsilon$ contrary to the assumption. □

The second lemma presents a folkloric renorming result.

Lemma 3.11. *Given a separable Banach space $(Y, \|\cdot\|)$ and an $\varepsilon > 0$, there is an equivalent Gâteaux smooth norm $|\cdot|$ on Y such that $|y| \leq \|y\| \leq (1+\varepsilon)|y|$ for every $y \in Y$.*

Proof. Let $\{y_1, y_2, \ldots\}$ be a countable dense subset of the unit ball B_Y. Put

$$|y^*| := \left(\|y^*\|^2 + \varepsilon \sum_{n=1}^{\infty} 2^{-n} \langle y^*, y_n \rangle^2 \right)^{1/2}, \quad y^* \in Y^*.$$

A folkloric (but not completely trivial) argument guarantees that $|\cdot|$ thus defined is an *equivalent dual norm* on Y and that we have

$$\|y^*\| \leq |y^*| \leq (1+\varepsilon)\|y^*\| \quad \text{for every } y^* \in Y^*.$$

Hence, denoting by the same symbol $|\cdot|$ the corresponding predual norm on Y, we have

$$|y| \leq \|y\| \leq (1+\varepsilon)|y| \quad \text{for every } y \in Y.$$

[2]That is, for any $L \in \mathcal{S}(X)$ and $M \in \mathcal{S}(Y)$ there is a pair $(L', M') \in \mathcal{E}$ such that $L \subset L'$ and $M \subset M'$.

Assume that $|\cdot|$ is not Gâteaux differentiable at some $0 \neq y \in Y$. Then there are distinct $y_1^*, y_2^* \in Y^*$, of norm 1, such that $\langle y_1^*, y\rangle = \langle y_2^*, y\rangle = |y|$; then $|y_1^* + y_2^*| = 2$. It follows that

$$2|y_1^*|^2 + 2|y_2^*|^2 - |y_1^* + y_2^*|^2 = 0,$$

and a convexity argument yields that $\langle y_1^* - y_2^*, y_n\rangle^2 = 0$ for all $n \in \mathbb{N}$. Thus $y_1^* = y_2^*$, a contradiction. $\square$

Proof of Theorem 3.9. If both X and Y are separable spaces, then the result follows from the first parts of Theorem 3.4, Proposition 3.7, and Lemma 3.11.

Thus we only need to show that the family of all pairs

$$(L, M) \in \mathcal{S}(X) \times \mathcal{S}(Y)$$

(containing $(\overline{x}, F(\overline{x}))$), such that $F(L) \subset M$ and that the mapping

$$F\restriction_L : L \to M$$

satisfies conditions (a) and (b) in Theorem 3.8 if $X, Y, K := \operatorname{dom} F - \overline{x}$ and F are replaced by L, M, $K \cap L$ and $F\restriction_L$, respectively, is cofinal in $\mathcal{S}(X) \times \mathcal{S}(Y)$. Indeed, as L and M are separable, there are Gâteaux smooth norms in L and M arbitrarily close to the given norms. Therefore, by Theorem 3.4 for any $\varepsilon > 0$ we have $\operatorname{sur} F\restriction_L (\overline{x}) \geq \rho - \varepsilon$, hence $\operatorname{sur} F\restriction_L (\overline{x}) \geq \rho$. Applying Lemma 3.10, we get the result.

In proving the cofinality of the family described in the previous paragraph, we can harmlessly assume that $\overline{x} = 0$. Given some $L_0 \in \mathcal{S}(X)$ and $M_0 \in \mathcal{S}(Y)$ (no loss of generality occurs if we assume that they contain, respectively, $\overline{x}$ and $F(\overline{x})$), we shall construct a sequence of pairs

$$(L_i, M_i) \in \mathcal{S}(X) \times \mathcal{S}(Y)$$

such that

(i) $L_0 \subset L_1 \subset L_2 \subset \cdots$ and $M_0 \subset M_1 \subset M_2 \subset \cdots$;
(ii) $M_n \subset T(L_n) \subset M_{n+1}$ and $rB_{M_n} \subset \operatorname{cl}\big(T(B_{L_n} \cap K)\big)$ for all $T \in \mathcal{T}$;
(iii) $F(L_n) = F(L_n \cap K) \subset M_{n+1}$.

If such sequences are found, we define L and M as the closures of $\bigcup L_n$ and $\bigcup M_n$, respectively. Clearly, both subspaces are separable. From (ii) we conclude that for any $T \in \mathcal{T}$ the closure of $T(L)$ coincides with M and, moreover, $rB_M \subset \operatorname{cl}(T(B_L \cap K))$. Applying now the density theorem (Theorem 2.4) we conclude that $r\mathring{B}_M \subset T(B_L \cap K)$. Finally, (iii) shows that $F\restriction_L$ is a mapping from $L \cap K$ into M. This concludes verification of (a) and (b) for L, M, $K \cap L$ and $F\restriction_L$.

The construction of the sequences (M_n) and (L_n) is not very complicated. Suppose we have already L_n and M_n. We first define M_{n+1} as the subspace spanned by the union of M_n, $F(L_n)$ and all $T(L_n)$, $T \in \mathcal{T}$ (profiting from the separability of $\mathcal{T}$). Then, once M_{n+1} and $\mathcal{T}$ are separable, we can easily find a separable subspace $L_n \subset L_{n+1} \subset X$ such that

$$rB_{M_{n+1}} \subset \operatorname{cl}\big(T(B_{L_{n+1}} \cap K)\big). \qquad \square$$

Based on the theorem, we can prove a far reaching extension of Clarke's inverse function theorem.

Theorem 3.12. *We posit the assumption of Theorem* 3.9 *and assume in addition that every* $T \in \mathcal{T}$ *is an invertible operator. Then* F^{-1} *has a graphical localization around* $F(\overline{x})$ *which is both single valued and Lipschitz continuous with constant* $1/\rho$*. In other words, there are neighborhoods* U *of* $\overline{x}$ *and* V *of* $\overline{y} = F(\overline{x})$ *and a* $(1/\rho)$*-Lipschitz mapping* $G : V \to X$ *such that* $G(\overline{y}) = \overline{x}$ *and* $F^{-1}(y) \cap U = \{G(y)\}$ *for all* $y \in V$*.*

Proof. As $\operatorname{sur} F(\overline{x}) \geq \rho$, there are $\delta > 0$ and $r > 0$ such that

$$B(F(\overline{x}), \delta) \subset F(B(\overline{x}, r) \cap \operatorname{dom} F).$$

Take some $y, y' \in B(F(\overline{x}), \delta)$, and choose x and x' in $B(\bar{x}, r) \cap \operatorname{dom} F$ such that $y = F(x)$ and $y' = F(x')$. By (3.13) there is a $T \in \mathcal{T}$ such that

$$\|F(x') - F(x) - T(x' - x)\| \leq \beta \|x' - x\|.$$

Then (3.14), along with invertibility of T, implies that $(\beta+\rho)T^{-1}(B_Y) \subset B_X$ which means that $\|T^{-1}\| \leq (\rho + \beta)^{-1}$. Therefore,

$$\begin{aligned}\|x' - x\| &\leq \|T^{-1}\| \, \|T(x' - x)\| \\ &\leq \frac{1}{\rho + \beta}\Big(\|y' - y - T(x' - x)\| + \|y' - y\|\Big) \\ &\leq \frac{\beta}{\beta + \rho}\|x' - x\| + \frac{1}{\beta + \rho}\|y' - y\|.\end{aligned}$$

Multiplying the latter inequality by $(\beta + \rho)/\rho$ and then rearranging it a bit we get that $\|x' - x\| \leq \rho^{-1}\|y' - y\|$. Taking $y' = y$ here, we get that the mapping

$$B\big(F(\overline{x}), \delta\big) \ni y \longmapsto F^{-1}(y) \cap B(\overline{x}, r)$$

is a single-valued graphical localization of F^{-1} around $F(\overline{x})$ for $\overline{x}$. Using the last estimate for distinct y and y', we get that this graphical localization is Lipschitz continuous with the constant $1/\rho$. □

As in the preceding subsection we can replace the compactness assumption by a suitable estimate for the measure of noncompactness of $\mathcal{T}$, that is, the following theorem is true.

Theorem 3.13. *We posit the assumption of Theorem* 3.9 *except that instead of assuming that* $\mathcal{T}$ *is relatively compact in* $\mathcal{L}(X, Y)$*, we suppose that* $\chi(\mathcal{T}) < \infty$*. Then* $\operatorname{sur} F(\overline{x}) \geq (\rho - \chi(\mathcal{T}))^+$*.*

The results presented in this subsection also have a long history starting with the famous theorem of Graves [18] which exactly corresponds to Theorem 3.9 with $\mathcal{T}$ being a singleton and $\overline{x} \in \operatorname{int}(\operatorname{dom} F)$. Together with the famous 1934 Lyusternik theorem, the result of Graves was the starting point for the development of the regularity theory of variational analysis. Theorem 3.8 was proved in [15] under a de facto assumption that $\overline{x}$ is in the interior of $\operatorname{dom} F$. An earlier version of the theorem under somewhat stronger assumptions was proved in [9].

Theorem 3.13 (with $\beta = 0$ and also $\overline{x}$ in the interior of $\operatorname{dom} F$) was proved by Páles in [27]. Finally, the main result of the subsection, Theorem 3.9, is a slightly more precise version of a recent result of the first two authors [6]. (We assume that (3.13) is satisfied only for $x, x' \in \operatorname{dom} F$ sufficiently close to $\overline{x}$, not for all elements of $\operatorname{dom} F$. This allows to exclude the influence of the size of $\operatorname{dom} F$ on $\operatorname{sur} F$.) In both [27] and [6], the proofs use a topological machinery including Michael's selection theorem and Schauder's fixed point theorem. Our proof, on the contrary, is purely analytic and is essentially based on the techniques of regularity theory of variational analysis and a renorming. We emphasize the role of separable reduction that allows to prove the meaningful part of the statement only for separable spaces.

Note, however, that the proofs given in [27] and [6] can be substantially simplified and shortened if the regularity criterion of Theorem 2.3 is incorporated into the argument.

Alternative proof of Theorem 3.9. With no loss of generality, we assume that $\overline{x} = 0$ and $F(\overline{x}) = 0$. We can also safely assume that $\mathcal{T}$ is a compact set: (3.14) is still valid for all elements of the closure of $\mathcal{T}$.

To begin with, take a small $\delta > 0$ that will be fixed throughout most of the proof. The obvious inclusion $(1-\delta)B_X - x \subset B_X$ if $\|x\| < \delta$ implies that for any $\lambda \in (0,1)$ and $x \in \delta\lambda B_X$,

$$(\lambda B_X) \cap (\operatorname{dom} F - x) \supset \big[((1-\delta)\lambda B_X) \cap \operatorname{dom} F\big] - x. \tag{3.15}$$

Set $\tau = \sup\{\|T\| : T \in \mathcal{T}\}$ and $N = \rho + \beta + \tau$. If $\lambda \le \varepsilon$, then for any $T \in \mathcal{T}$ and $x \in \delta\lambda B_X$ we get by (3.15) and (3.14)

$$\begin{aligned} T\big((\lambda B_X) \cap (\operatorname{dom} F - x)\big) &\supset (1-\delta)T\big((\lambda B_X) \cap \operatorname{dom} F\big) - Tx \\ &\supset \lambda\big[\rho + \beta - N\delta\big]B_Y. \end{aligned} \tag{3.16}$$

We assume in what follows that $\lambda < \varepsilon$ is such that (3.13) holds for all $x, x' \in (\delta\lambda B_X) \cap \operatorname{dom} F$. For a time being, fix an $x \in \delta\lambda B_X$.

Fix for a longer while a nonzero $v \in \varepsilon(\rho + \beta - N\delta)B_Y$. As follows from (3.16), for any $T \in \mathcal{T}$, there is an $h \in (\varepsilon B_X) \cap (\operatorname{dom} F - x)$ such that $Th = v$ and $\|h\| \le (\rho + \beta - N\delta)^{-1}\|v\|$. For any $T \in \mathcal{T}$, let us fix such an $h =: h(T)$ and further put

$$U(T) := \big\{T' \in \mathcal{T} : \|T'h(T) - v\| < \delta\|v\|\big\}.$$

This set is nonempty as $T \in U(T)$, it is obviously open and the union of such sets when T ranges through $\mathcal{T}$ covers $\mathcal{T}$. As $\mathcal{T}$ is compact, we can choose a finite subcovering

$$\{U(T_1), \ldots, U(T_k)\}$$

of it. Set $h_i := h(T_i)$, and let $\{\alpha_1(\cdot), \ldots, \alpha_k(\cdot)\}$ be a continuous partition of unity subordinated to this subcovering; that is, $\alpha_i(\cdot) \ge 0$, $\alpha_i(T) = 0$ if $T \in \mathcal{T} \setminus U(T_i)$, and $\alpha_1(\cdot) + \cdots + \alpha_i(\cdot) = 1$. (For instance, we can take $\beta_i(T)$

equal to the distance from T to the complement of $U(T_i)$ in $\mathcal{T}$ if $T \in U(T_i)$ (and zero otherwise) and set $\alpha_i(T) = \beta_i(T)/(\sum_i \beta_i(T))$.)

Next we define a mapping $\tilde{h} : \mathcal{T} \to X$ by

$$\tilde{h}(T) = \sum_{i=1}^{k} \alpha_i(T) h_i, \quad T \in \mathcal{T}.$$

Clearly, $\tilde{h}$ is a continuous mapping, $\operatorname{range} \tilde{h} \subset \operatorname{dom} F$ (as $\operatorname{dom} F$ is convex). Furthermore, for each $T \in \mathcal{T}$ we have

$$\|\tilde{h}(T)\| \le (\rho + \beta - N\delta)^{-1}\|v\| \quad \text{and} \quad \|T\tilde{h}(T) - v\| < \delta\|v\|. \tag{3.17}$$

The first inequality is immediate from the definition of $\tilde{h}(T)$. To verify the second, we observe that $\alpha_i(T) > 0$ only if $\|T(h(T_i))-v\| < \delta\|v\|$ and therefore

$$\|T(\tilde{h}(T)) - v\| = \left\|\sum \alpha_i(T)\big(Th_i - v\big)\right\| \le \sum \alpha_i(T)\|Th_i - v\| < \delta\|v\|,$$

and the second inequality in (3.17) follows. Note finally that the mapping $\tilde{h}(\cdot)$ depends on the choice of v, so it would be natural to denote it by $h_v(\cdot)$.

Next for any $h \in X$ let $T_x(h)$ be the collection of all $T \in \mathcal{T}$ such that (3.13) holds with $x' := x + h$. Then the set $T_x(h)$ is nonempty and convex if $x + h \in (\delta\lambda B_X) \cap \operatorname{dom} F$ and the graph of the set-valued mapping $h \mapsto T_x(h)$ is closed, the latter being true because $\operatorname{dom} F$ is a closed set by the assumption and F is continuous (even Lipschitz) on $\operatorname{dom} F$ by (3.13).

Finally, take any $x \in \operatorname{dom} F$ with $\|x\| < \delta\lambda$ and any $y \ne F(x)$. Put

$$v := -\xi(y - F(x)),$$

where ξ is small enough to guarantee that

$$\|v\| \le \lambda(\rho + \beta - N\delta) \quad \text{and} \quad \|x + h_v(T)\| < \delta\lambda \quad \text{for all } T.$$

Then $h_v(T)$ belongs to the domain of T_x for all $T \in \mathcal{T}$, that is,

$$(T_x \circ h_v)(T) = T_x(h_v(T)) \ne \emptyset$$

and as $h_v(\cdot)$ is continuous, the graph of $T_x \circ h_v$ is closed. This is a convex-valued mapping from $\mathcal{T}$ into itself and, as $\mathcal{T}$ is compact and convex, we can apply Gliksberg's extension of the Kakutani fixed point theorem [16, 17] (based on Brouwer's fixed point theorem) and conclude that there is a $\widehat{T} \in \mathcal{T}$ such that

$$T_x\big(h_v\big(\widehat{T}\big)\big) \ni \widehat{T}.$$

Set $\hat{h} := h_v(\widehat{T})$. By (3.17), we have

$$\begin{aligned}
\big\|y - F\big(x + \hat{h}\big)\big\| &\le \|y - F(x) - v\| + \big\|F\big(x + \hat{h}\big) - F(x) - \widehat{T}\hat{h}\big\| + \big\|\widehat{T}\hat{h} - v\big\| \\
&\le \|y - F(x)\| - \|v\| + \beta\big\|\hat{h}\big\| + \delta\|v\| \\
&\le \|y - F(x)\| - \left(1 - \left(\frac{\beta}{\rho + \beta - N\delta} + \delta\right)\right)\|v\| \\
&\le \|y - F(x)\| - \big(\rho - N\delta - (\rho + \beta - N\delta)\delta\big)\big\|\hat{h}\big\|.
\end{aligned}$$

In other words, for any $\rho' < \rho$ we can find (by choosing a suitable δ) a neighborhood of $\overline{x}$ such that for any $x \in \operatorname{dom} F$ in this neighborhood and any $y \neq F(x)$ there is an $h \neq 0$ such that $\|y - F(x+h)\| \le \|y - F(x)\| - \rho'\|h\|$. The proof is now completed by a reference to the regularity criterion of Theorem 2.3. □

4. Concluding remarks and comparison of the results

Theorems 3.2, 3.4 and 3.9 seem to represent the trio of the most advanced (for the moment) results with primal estimates for regularity rates. The important point is that the three theorems are different and neither of them implies the other, unless the spaces are finite dimensional. In this latter case, the estimates for the rate of surjection provided by both parts of Theorem 3.2 are never worse than the estimates of the other two theorems. The following example demonstrates the phenomenon.

Example 4.1. Let $X = \mathbb{R}^2$, with ℓ_∞-norm, $Y = \mathbb{R}$ and $f(x) = |x_1| - |x_2|$. Clearly $\operatorname{sur} f(x) = 1$ for any x. This function is directionally differentiable at every x and its contingent derivative (as a mapping from $\mathbb{R}^2$ into $\mathbb{R}$) coincides with its standard directional derivative

$$f'(x;h) = \lim_{t \downarrow 0} t^{-1}\big(f(x+th) - f(x)\big).$$

We have

$$Df(x;h) = f'(x;h) = \begin{cases} (\operatorname{sign} x_1)h_1 - (\operatorname{sign} x_2)h_2 & \text{if } x_1 \neq 0,\ x_2 \neq 0; \\ (\operatorname{sign} x_1)h_1 \pm h_2 & \text{if } x_1 \neq 0,\ x_2 = 0; \\ \pm h_1 - (\operatorname{sign} x_2)h_2 & \text{if } x_1 = 0,\ x_2 \neq 0; \\ \pm h_1 \pm h_2 & \text{if } x_1 = x_2 = 0. \end{cases}$$

Clearly, for any x we can find an h, either equal to $(\pm 1, 0)$ or to $(0, \pm 1)$ such that $f'(x;h)$ is equal to 1 or -1. In other words, the condition of Theorem 3.2 is satisfied with $\lambda = 0$ and $c = 1$ and the theorem gives the precise value of the rate of surjection.

On the other hand, the generalized gradient of f at zero is the ℓ_∞ unit ball in $\mathbb{R}^2$: $\{y : \max\{|y_1|, |y_2|\} \le 1\}$, so any collection $\mathcal{T}$ corresponding to $\beta < 1$ would contain zero. This means that Theorem 3.9 provides only the trivial estimate $\operatorname{sur} f(0) \ge 0$. Similar arguments lead to the same estimate in Theorem 3.4.

The point is that in a finite-dimensional situation Theorem 3.2 is a source result for an exact estimate of the rate of surjection (see [11]). On the other hand, the advantage of Theorem 3.4 and especially Theorem 3.9 in this case is due to greater simplicity of their application if the mapping is Lipschitz on its domain. Indeed, the simplest and the most natural set of operators in this case is the generalized Jacobian. To compute it, we only need to know derivatives of the mappings on some set of full measure. On the other hand, Theorem 3.2 requires calculation of $DF(x)$, or the tangent cone to the graph, or the Dini–Hadamard lower directional derivative at every point which may require much more effort. So if we only wish to know whether the mapping is open, Theorems 3.4 and 3.9 may be more convenient.

Things change in the infinite-dimensional case when the quality of approximation provided by the contingent cone may be very poor. For instance, it is an easy matter to verify that (unlike the second part of Theorem 3.2) both Theorems 3.4 and 3.9 work well in Example 3.3 with $\mathcal{T} = \{I\}$ (I being the identity mapping) and $\mathcal{H}(x) = \{x\}$ respectively, $\rho = 1/4$ and $\beta = 3/4$. Performance of Theorem 3.2 can be improved under additional assumptions that improve the behavior of the contingent cones (e.g., sleekness [3]).

Now about the relations between Theorems 3.4 and 3.9. On the one hand, the assumptions of the first theorem are noticeably weaker. The positively homogeneous mapping generated by a compact set of linear operators is necessarily bounded (its upper norm $\|\mathcal{H}\| = \sup\{\|y\| : y \in \mathcal{H}(x),\ \|x\| \le 1\}$ is necessarily finite) and moreover, it is Lipschitz with respect to the Hausdorff metric. None of these properties is required for $\mathcal{H}$ in Theorem 3.4. The most restrictive is the convexity assumption. It is essential (look for $f(x) = |x|$ and $\mathcal{T} = \{-1, 1\}$). But say, in Example 4.1 with $\overline{x} = 0$ every convex set of linear functions satisfying property (a) of Theorem 3.8 contains zero.

On the other hand, the obvious advantage of the last theorem is that it is valid in all Banach spaces while the first is not. The cause is the absence of a separable reduction for the property (3.6). Whether or not such a reduction is possible is still unclear.

In addition, working with sets of linear operators can be much more convenient. The problem is that there is no way known to find suitable sets of linear operators unless $\dim Y < \infty$ (see [28]). Existence theorems for linear selections of positively homogeneous set-valued mappings may offer a possible way to bypass this difficulty. A general condition that guarantees that a positively homogeneous mapping is generated by its linear selections is given in [26], but it does not seem easily verifiable in practical situations. Some more specialized conditions for concrete types of prederivatives can be found in [19]. Still it is not clear how to translate properties of the set-valued mapping into properties of its individual linear selections, unless both spaces are reflexive (see [15]).

Acknowledgments

The first author was supported by the project GA15-00735S. The second author was supported by grant P201/12/0290 and by RVO: 67985840.

References

[1] R. R. Akhmerov, M. I. Kamenskii, A. S. Potapova, A. E. Rodkina and B. N. Sadovskii, *Measure of Noncompactness and Condensing Operators*. Birkhüser, Basel, 1992.

[2] J.-P. Aubin, *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*. In: Mathematical

Analysis and Applications, Part A, Adv. in Math. Suppl. Stud. 7, Academic Press, New York, 1981, 159–229.

[3] J. P. Aubin and H. Frankowska, *Set-Valued Analysis.* Systems & Control: Foundations & Applications 2, Birkhäuser Boston, Boston, MA, 1990.

[4] D. Azé, *A unified theory for metric regularity of multifunctions.* J. Convex Anal. **13** (2006), 225–252.

[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] R. Cibulka and M. Fabian, *A note on Robinson-Ursescu and Lyusternik-Graves theorem.* Math. Program. **139** (2013), 89–101.

[7] E. De Giorgi, A. Marino and M. Tosques, *Problemi di evoluzione in spazi metrici e curve di massima pendenza.* Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **68** (1980), 180–187.

[8] R. Deville, G. Godefroy and V. Zizler, *Smoothness and Renormings in Banach Spaces.* Pitman Monographs and Surveys in Pure and Applied Mathematics 64, Longman Scientific & Technical, Harlow, 1993.

[9] P. H. Dien, *Some results on locally Lipschitzian mappings.* Acta Math. Vietnam. **6** (1981), 97–105.

[10] 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.

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

[12] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings.* Springer Monographs in Mathematics, Springer, Dordrecht, 2009.

[13] N. Dunford and J. T. Schwartz, *Linear Operators.* I. *General Theory.* Interscience Publishers, New York, 1958.

[14] M. Fabian, P. Habala, P. Hájek, V. Montesinos and V. Zizler, *Banach Space Theory: The Basis for Linear and Nonlinear Analysis.* CMS Books Math., Springer, New York, 2011.

[15] M. Fabian and D. Preiss, *A generalization of the interior mapping theorem of Clarke and Pourciau.* Comment. Math. Univ. Carolin. **28** (1987), 311–324.

[16] I. L. Glicksberg, *A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points.* Proc. Amer. Math. Soc. **3** (1952), 170–174.

[17] A. Granas and J. Dugunji, *Fixed Point Theory.* Springer, New York, 2003.

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

[19] A. D. Ioffe, *Nonsmooth Analysis: Differential calculus of nondifferentiable mappings.* Trans. Amer. Math. Soc. **266** (1981), 1–56.

[20] A. D. Ioffe, *On the local surjection property.* Nonlinear Anal. **11** (1987), 565–592.

[21] A. D. Ioffe, *Metric regularity and subdifferential calculus.* Russian Math. Surveys **55** (2000), 501–558.

[22] A. D. Ioffe, *Regularity on a fixed set.* SIAM J. Optim. **21** (2011), 1345–1370.

[23] A. D. Ioffe, *Separable reduction of metric regularity properties.* In: Constructive Nonsmooth Analysis and Related Topics, V. F. Demyanov, P. M. Pardalos and M. Batsin, eds., Springer Optim. Appl. 87, Springer, New York, 2013, 25–37.

[24] A. D. Ioffe, *Metric regularity, fixed points and some associated problems of variational analysis.* J. Fixed Point Theory Appl. **15** (2014), 67–99.

[25] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation.* I. Grundlehren Math. Wiss. 300, Springer, Berlin, 2006.

[26] Z. Páles, *Linear selections for set-valued functions and extensions of bilinear forms.* Arch. Math. (Basel) **62** (1994), 427–432.

[27] Z. Páles, *Inverse and implicit function theorems for nonsmooth maps in Banach spaces.* J. Math. Anal. Appl. **209** (1997), 202–220.

[28] Z. Páles and V. Zeidan, *Infinite dimensional Clarke generalized Jacobian.* J. Convex Anal. **14** (2007), 433–454.

[29] J.-P. Penot, *Metric regularity, openness and Lipschitzian behavior of multi-functions.* Nonlinear Anal. **13** (1989), 629–643.

[30] J.-P. Penot, *Calculus without Derivatives.* Grad. Texts in Math. 266, Springer, New York, 2013.

[31] V. Pták, *A quantitative refinement of the closed graph theorem.* Czechoslovak Math. J. **24** (1974), 503–506.

[32] K. Sh. Tsiskaridze, *Extremal problems in Banach spaces.* In: Nekotorye Voprosy Matematicheskoy Theorii Optimalnogo Upravleniya (Some Problems of the Mathematical Theory of Optimal Control), Inst. Appl. Math., Tbilisi State Univ. 1975 (in Russian).

Radek Cibulka
NTIS – New Technologies for the Information Society
Department of Mathematics
Faculty of Applied Sciences
University of West Bohemia
Univerzitní 22
306 14 Pilsen
Czech Republic
e-mail: `cibi@kma.zcu.cz`

Marián Fabian
Mathematical Institute of Czech Academy of Sciences
Žitná 25
115 67 Praha 1
Czech Republic
e-mail: `fabian@math.cas.cz`

Alexander D. Ioffe
Department of Mathematics
Technion - Israel Institute of Technology
Haifa 32000
Israel
e-mail: `ioffe@math.technion.ac.il`