ON A SET-COVERING PROPERTY OF MULTIVALUED MAPPINGS

AMOS UDERZO

To the memory of Aleksander Moiseevich Rubinov (1940-2006)

ABSTRACT. In this paper, a set-oriented variant for multivalued mappings of the well-known property of openness at a linear rate is studied. Among other examples, a simply characterized class of closed convex processes between Banach spaces, which satisfies such a covering behaviour, is singled out. Equivalent reformulations of this property and its stability under Lipschitz perturbations are investigated in a metric space setting. Applications to the solvability of set-valued inclusions and to the exact penalization of optimization problems with set-inclusion constraints are discussed.

1. INTRODUCTION AND PRELIMINARIES

The property of being surjective describes an elementary set-theoretic behaviour of mappings which, in synergy with specific (topological, metric, linear, and so on) structures on the domain and on the range sets, may afford valuable consequences. A paradigma of this phenomenon can be seen in the celebrated Banach-Schauder open mapping theorem for linear operators: in a context, where the metric completeness interacts with linearity and continuity, the property of a mapping of being onto turns out to imply its openness (namely, images of open sets remain open). Even though some quantitative estimates did appear already in the original statement of this result (see, for instance, Theorem 10 in Ch. X of [6]), the merely topological formulation of it left somehow hidden certain metric aspects of this "openness preservation law". The potential of them was understood several decades later, only after the open mapping theorem was extended in a local form to nonlinear mappings by L.A. Lyusternik (see [22]) and L.M. Graves (see [17]). Their far-reaching extensions, indeed, paved the way to enlightening the interconnections of the quantitative surjective behaviour of a mapping with the Lipschitzian behaviour of its inverse and with error bounds for the solution set to the related generalized equations (relationships between such "regularity phenomena" were observed by S.M. Robinson in [30] and A.A. Milyutin in [10], and later proved also in [23, 27]), and, consequently, to explore links with the metric fixed point theory (see [1, 12, 19]). This led to distil

the notion of metric regularity as an essential tool for the analysis of various stability and sensitivity issues in modern variational analysis, optimization and control theory (see detailed historical comments in [12, 18, 25, 32]). As a consequence, the covering behaviour of single as well as of set-valued mappings has been the main subject of many investigations (see, among the others, [1, 7, 10, 15, 24, 26, 35]). In them, depending on possible applications or on specific issues of the related theory to be investigated, several variations of the concept of covering behaviour itself have been considered: for instance, a reader will find openness at a linear or at a more general rate, local covering, global covering, openness restricted to given sets, point based openness, linear semiopenness, and so on.

In the present paper, a set-oriented variant of the notion of openness at a linear rate in a metric space setting is considered, which applies only to multivalued mappings. Roughly speaking, such a property postulates that the whole enlargement of images through a given mapping are covered by the image of a single element near the reference point, instead of by the image of an entire ball around it. Such a requirement clearly imposes severe restrictions on the covering behaviour of a set-valued mapping, yet it happens to be fulfilled in various contexts, which are relevant to variational analysis and optimization. Furthermore, as shown in the present study, it exhibits nice robustness features in the presence of various types of perturbations. Other motivations for the interest in this set-covering behaviour come from set-inclusion problems, namely generalized equations where the inclusion of a single element into images of a given multifunction is replaced by the inclusion of an entire set. With respect to such kind of problems, it seems that solvability and solution stability can be hardly approached as far as working with conventional covering notions.

The contents of the paper are organized as follows. In the rest of the current section, basic notations, preliminary notions and related facts, that will be employed throughout the paper, are recalled. In Section 2 the main set-covering property under study is introduced. Several contexts in which it emerges are discussed. In particular, the class of closed convex processes having this behaviour is singled out. Then, conditions for such a property to hold are established in a metric space setting. Its stability under under some kinds of perturbation is also considered. Section 3 is devoted to explore some applications to the existence of set-inclusion points, with related error bound estimates, and to the exact penalization of constrained optimization problems.

Whenever x is an element of a metric space (X, d) and r is a positive real, $\mathrm{B}(x, r) = \{z \in X : d(z, x) \leq r\}$ denotes the closed ball with center x and radius r. By $\mathrm{dist}\,(x, S) = \inf_{z \in S} d(z, x)$ the distance of x from a subset $S \subseteq X$ is denoted, with the convention that $\mathrm{dist}\,(x, \varnothing) = +\infty$. The r-enlargement of a set $S \subseteq X$ is indicated by $\mathrm{B}(S, r) = \{x \in X : \mathrm{dist}\,(x, S) \leq r\}$. Given sets $A, B \subseteq X$, the excess of A over B is indicated by $\mathrm{exc}\,(A, B) = \sup_{a \in A} \mathrm{dist}\,(a, B)$, while the Hausdorff distance of A and B by $\mathrm{Haus}(A, B) = \max\{\mathrm{exc}\,(A, B), \mathrm{exc}\,(B, A)\}$. Recall that a set-valued mapping $\Phi : X \rightrightarrows Y$ between metric spaces is said to be Lipschitz on X with constant $l \geq 0$ provided

$$\mathrm{exc}\,(\Phi(x_1), \Phi(x_2)) \leq l d(x_1, x_2), \quad \forall x_1, x_2 \in X.$$

If the above inequality is satisfied in a neighbourhood of a given point $\bar{x} \in X$, Φ is said to be locally Lipschitz around $\bar{x}$. A set-valued mapping $\Psi : X \rightrightarrows Y$ between metric spaces is called Hausdorff upper semicontinuous (henceforth, u.s.c.) at $x_0 \in X$ is for every $\epsilon > 0$ there exists $\delta > 0$ such that $\Psi(x) \subseteq \mathrm{B}(\Psi(x_0), \epsilon)$, for every $x \in \mathrm{B}(x_0, \delta)$. The domain and the graph of $\Psi : X \rightrightarrows Y$ are denoted by $\mathrm{dom}\, \Psi$ and $\mathrm{grph}(\Psi)$, respectively. Throughout the paper, any mapping $\Psi : X \rightrightarrows Y$ will be assumed to have $\mathrm{dom}\, \Psi = X$ and to take closed values, unless otherwise stated. In any vector space, the null element is marked by $\mathbf{0}$, and the related notations $\mathbb{B} = \mathrm{B}(\mathbf{0}, 1)$ and $\mathbb{S} = \mathrm{bd}\, \mathbb{B} = \mathbb{B} \backslash \mathrm{int}\, \mathbb{B}$ are adopted, where int and bd indicate the topological interior and boundary of a given set, respectively.

Remark 1.1. In the sequel, the following consequence of the Lipschitz property of a set-valued mapping on its excess function will be used: if Φ is Lipschitz on X with constant l, then for any nonempty set $S \subseteq Y$, the function $x \mapsto \mathrm{exc}\,(\Phi(x), S)$ is Lipschitz on X with the same constant l. Indeed, for every $x_1, x_2 \in X$ it is true that

$$
\begin{aligned}
\mathrm{exc}\,(\Phi(x_2), S) &= \sup_{y \in \Phi(x_2)} \mathrm{dist}\,(y, S) \le \sup_{y \in \mathrm{B}(\Phi(x_1), ld(x_1, x_2))} \mathrm{dist}\,(y, S) \\
&\le \sup_{y \in \mathrm{B}(\Phi(x_1), ld(x_1, x_2))} \mathrm{dist}\,(y, \Phi(x_1)) + \mathrm{exc}\,(\Phi(x_1), S) \\
&\le ld(x_1, x_2) + \mathrm{exc}\,(\Phi(x_1), S).
\end{aligned}
$$

Another property of the excess function associated with a pair of set-valued mappings, that will be used in the sequel, is stated next.

Lemma 1.2. *Let $\Psi : X \rightrightarrows Y$ and $\Phi : X \rightrightarrows Y$ be given set-valued mappings between metric spaces. Suppose that:*
  (i) *Ψ is Hausdorff u.s.c. at $x_0 \in X$;*
  (ii) *Φ is Lipschitz on X, with constant $l \ge 0$.*
Then, the function $\mathrm{exc}_{\Phi, \Psi} : X \longrightarrow [0, +\infty]$ defined as

$$
\mathrm{exc}_{\Phi, \Psi}(x) = \mathrm{exc}\,(\Phi(x), \Psi(x)), \quad x \in X,
$$

is lower semicontinuous (for short, l.s.c.) at $x_0 \in X$.

Proof. Since $\mathrm{exc}_{\Phi, \Psi}$ acts on a metric space, it suffices to show that for every sequence $(x_n)_{n \in \mathbb{N}}$ in X, with $x_n \to x_0$ as $n \to \infty$, it results in

$$
(1.1) \qquad\qquad \mathrm{exc}_{\Phi, \Psi}(x_0) \le \liminf_{n \to \infty} \mathrm{exc}_{\Phi, \Psi}(x_n).
$$

It is immediate to observe that

$$
\begin{aligned}
\mathrm{exc}_{\Phi, \Psi}(x_0) &\le \mathrm{exc}\,(\Phi(x_0), \Phi(x_n)) + \mathrm{exc}_{\Phi, \Psi}(x_n) + \mathrm{exc}\,(\Psi(x_n), \Psi(x_0)) \\
&\le ld(x_0, x_n) + \mathrm{exc}_{\Phi, \Psi}(x_n) + \mathrm{exc}\,(\Psi(x_n), \Psi(x_0)),
\end{aligned}
$$

where the second inequality follows from hypothesis (ii). Since Ψ is is Hausdorff u.s.c. at x_0, then $\lim_{n \to \infty} \mathrm{exc}\,(\Psi(x_n), \Psi(x_0)) = 0$. Thus, one obtains

$$
\begin{aligned}
\mathrm{exc}_{\Phi, \Psi}(x_0) &\le \liminf_{n \to \infty} [ld(x_0, x_n) + \mathrm{exc}_{\Phi, \Psi}(x_n) + \mathrm{exc}\,(\Psi(x_n), \Psi(x_0))] \\
&\le \lim_{n \to \infty} ld(x_0, x_n) + \liminf_{n \to \infty} \mathrm{exc}_{\Phi, \Psi}(x_n)
\end{aligned}
$$

$$+ \lim_{n\to\infty} \mathrm{exc}\,(\Psi(x_n), \Psi(x_0)) = \liminf_{n\to\infty} \mathrm{exc}_{\Phi,\Psi}(x_n).$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2. Set-covering mappings

2.1. **The main property and examples.** By a global covering behaviour of a multifunction $\Psi : X \rightrightarrows Y$ acting between metric spaces the following property is usually meant: there exists a constant $\alpha > 0$ such that

(2.1) $$\Psi(\mathrm{B}(x, r)) \supseteq \mathrm{B}(\Psi(x), \alpha r), \quad \forall x \in X, \ \forall r > 0$$

(see, for instance, [1, 2, 9, 16]).

The main notion here under study comes up as a set-oriented variant of the above property, as stated below.

Definition 2.1. A set-valued mapping $\Psi : X \rightrightarrows Y$ between metric spaces is said to be *set-covering* on X with constant α if there exists a positive real α such that

(2.2) $$\forall x \in X, \ \forall r > 0 \ \exists u \in \mathrm{B}(x, r) \text{ such that } \Psi(u) \supseteq \mathrm{B}(\Psi(x), \alpha r).$$

From Definition 2.1 it is clear that, if a set-valued mapping is set-covering with constant α, then it is covering in the sense of (2.1), with the same constant. The converse is not true, as readily illustrated in the counterexamples below. Other immediate consequences of Definition 2.1 are the facts that Ψ is onto and that densely on X it takes values with nonempty interior.

Example 2.2. Let $X = \mathbb{R}$ and $Y = \mathbb{R}^2$ be endowed with their usual (Euclidean) metric structure. Consider the set-valued mapping $\Psi : \mathbb{R} \rightrightarrows \mathbb{R}^2$ given by $\Psi(x) = |x|\mathbb{S}$. It is not difficult to see that Ψ is covering on $\mathbb{R}$ with constant $\alpha = 1$, whereas it fails to fulfil Definition 2.1 for any $\alpha > 0$, as it is $\mathrm{int}\,\Psi(x) = \varnothing$ for every $x \in \mathbb{R}$. Again, the mapping $\Psi : \mathbb{R} \rightrightarrows \mathbb{R}^2$ given by $\Psi(x) = \mathrm{B}(x, 1)$ is covering with constant 1, but is not set-covering, even if it takes images with nonempty interior. This second mapping shows that, while Definition 2.1 forces a mapping to have images with nonempty interior in a dense subset of X, this topological requirement is only necessary.

Example 2.3. Let $\delta : X \longrightarrow [0, +\infty)$ be a function defined on a metric space (X, d) and satisfying the condition

(2.3) $$\widehat{\alpha} := \inf_{x \in X} \inf_{r > 0} \sup_{u \in \mathrm{bd}\,\mathrm{B}(x,r)} \frac{\delta(u) - \delta(x)}{d(u, x)} > 0,$$

and let (Y, d) be a metric space. For any fixed $y_0 \in Y$, the set-valued mapping $\Psi : X \rightrightarrows Y$, defined by

$$\Psi(x) = \mathrm{B}(y_0, \delta(x)),$$

is set-covering on X with any constant $\alpha \in (0, \widehat{\alpha})$. Indeed, fixed such an α, let $x \in X$ and $r > 0$. Take $\epsilon > 0$ in such a way that $\alpha + 2\epsilon < \widehat{\alpha}$. By condition (2.3), corresponding to ϵ, there exists $u \in \mathrm{bd}\,\mathrm{B}(x, r)$ such that $\delta(u) \geq \delta(x) + (\widehat{\alpha} - \epsilon)r$. Thus, if $y \in \mathrm{B}(\Psi(x), \alpha r)$, that is $\mathrm{dist}\,(y, \mathrm{B}(y_0, \delta(x))) \leq \alpha r$, it results in

$$d(y, y_0) < \delta(x) + (\alpha + \epsilon)r < \delta(x) + (\widehat{\alpha} - \epsilon)r \leq \delta(u).$$

This means that $y \in B(y_0, \delta(u)) = \Psi(u)$. Since $u \in B(x, r)$, the requirement (2.2) is fulfilled. Notice that, whenever X is, in particular, a normed space, taking $\delta(\cdot) = \|\cdot\|$, one finds $\widehat{\alpha} = 1$, so condition (2.3) is valid.

Remark 2.4. It is worth mentioning that, whereas the covering behaviour for multifunctions defined by (2.1) can be regarded as a direct generalization to the set-valued context of a similar behaviour for single-valued mappings, this can not be said for the notion of set-covering. In fact, single-valued mappings, considered as a multivalued one, even though covering can never fulfil such an inclusion as (2.2). This should stress the difference between set-covering and the standard notion of covering.

Below some natural circumstances, in which the covering behaviour formalized in Definition 2.1 appear, are presented.

Example 2.5. (Solution mappings to systems of sublinear inequalities) Let $X = \mathbb{R}^n$ be metrized with the norm $\|\cdot\|_\infty$ and let $Y = \mathbb{R}^m$ be endowed with its usual (Euclidean) metric structure. Suppose that n functions $p_i : \mathbb{R}^m \longrightarrow \mathbb{R}$, with $i = 1, \ldots, n$, are given, which are sublinear on $\mathbb{R}^m$, i.e. such that

$$p_i(\mathbf{0}) = 0, \quad p_i(ty) = tp_i(y), \quad \forall t > 0, \forall y \in \mathbb{R}^m, \quad \text{and } p_i \text{ convex on } \mathbb{R}^m.$$

Set

$$\|p_i\|_* = \max\{\|y^*\| : \ y^* \in \partial p_i(\mathbf{0})\}, \quad i = 1, \ldots, n,$$

where $\partial p_i(\mathbf{0})$ denotes the subdifferential of p_i at $\mathbf{0}$ in the sense of convex analysis, and

$$\|p\|_* = \max_{i=1,\ldots,n} \|p_i\|_*.$$

Notice that $\|p_i\|_*$ and $\|p\|_*$ are well defined and finite, as each $\partial p_i(\mathbf{0})$ is a nonempty compact subset of $\mathbb{R}^m$. In what follows, it is assumed that $\|p\|_* > 0$, as the case $\|p\|_* = 0$ leads to $p_i \equiv 0$ for every $i = 1, \ldots, n$, which is of minor interest here. Consider the solution mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ associated with a parameterized inequality system involving functions p_i as follows

$$\Psi(x) = \{y \in \mathbb{R}^m : \ p_i(y) \leq |x_i|, \quad \forall i = 1, \ldots, n\}.$$

Ψ clearly takes nonempty closed and convex values. Let us show that Ψ is set-covering, with constant $\alpha = 1/\|p\|_*$. To do so, fixed $x \in \mathbb{R}^n$ and $r > 0$, take $y \in B(\Psi(x), r/\|p\|_*)$. This implies the existence of $v \in \Psi(x)$ such that $\|y-v\| \leq r/\|p\|_*$. Thus, it must be

$$p_i(v) \leq |x_i|, \quad \forall i = 1, \ldots, n.$$

By the sublinearity of each p_i, one has

$$p_i(y) \leq p_i(y - v) + p_i(v) \leq \|p\|_* \|y - v\| + |x_i| \leq r + |x_i|, \ \forall i = 1, \ldots, n.$$

Consequently, by defining $u \in \mathbb{R}^n$ as

$$u_i = \begin{cases} r + x_i, & \text{if } x_i \geq 0, \\ -r + x_i, & \text{if } x_i < 0, \end{cases}$$

it results in

$$p_i(y) \leq |u_i|, \quad \forall i = 1, \ldots, n,$$

and hence

$$\Psi(u) \supseteq B\left(\Psi(x), \frac{r}{\|p\|_*}\right),$$

with $\|u - x\|_\infty \le r$. If $\mathbb{R}^n$ is remetrized through another (equivalent) norm, Ψ still remains set-covering, but with a different constant.

Example 2.6. (Epigraphical mappings in partially ordered normed spaces) Let $(\mathbb{X}, \|\cdot\|)$ and $(\mathbb{Y}, \|\cdot\|)$ be real normed spaces. Suppose that on $\mathbb{Y}$ a partial order relation $\le_\mathbb{Y}$ is induced by a closed, convex, pointed cone $\mathbb{Y}_+ \subseteq \mathbb{Y}$, in the sense that $y_1 \le_\mathbb{Y} y_2$ iff $y_2 - y_1 \in \mathbb{Y}_+$. Let us introduce the following assumption on the interplay bewteen the partial order and the metric structure on $\mathbb{Y}$:

$$(2.4) \qquad \exists \gamma \in [1, +\infty) : \ \forall r > 0 \ \exists \check{y} \in \mathbb{Y} : \ \check{y} \le_\mathbb{Y} y, \ \forall y \in r\mathbb{B} \ \text{ and } \ \|\check{y}\| = \gamma r.$$

Of course, because of the linearity of the partial order, assumption (2.4) applies also to balls with center at each point of $\mathbb{Y}$. Such an assumption is verified, for instance, if $\mathbb{Y} = \mathbb{R}^m$, $\mathbb{Y}_+ = \mathbb{R}^m_+$, and $\|\cdot\|_p$ is a p-norm, with $p \in [1, +\infty)$, by the constant $\gamma = m^{1/p}$. Otherwise, if $\mathbb{Y} = \mathbb{R}^m$ is normed by $\|\cdot\|_\infty$, one has $\gamma = 1$. If $\mathbb{Y} = C([0,1])$, $\mathbb{Y}_+ = \{x \in C([0,1]) : \ x(t) \ge 0, \quad \forall t \in [0,1]\}$ and the metric structure on $\mathbb{Y}$ is induced by the norm $\|\cdot\|_\infty$, then again assumption (2.4) is true with $\gamma = 1$. Instead, if the space $(\ell_p(\mathbb{N}), \|\cdot\|_p)$, with $p \in [1, +\infty)$ is partially ordered by the componentwise order relation, assumption (2.4) fails to be verified. If $(\ell_\infty(\mathbb{N}), \|\cdot\|_\infty)$ is partially ordered by the same order relation, the above assumption is verified with $\gamma = 1$.

Now, let $f : \mathbb{X} \longrightarrow \mathbb{Y}$ be a mapping covering on $\mathbb{X}$ with constant $\alpha > 0$, i.e. such that

$$(2.5) \qquad\qquad f(B(x, r)) \supseteq B(f(x), \alpha r), \quad \forall x \in \mathbb{X}, \ \forall r > 0.$$

It is possible to show that its epigraphical set-valued mapping $\text{epi}_f : \mathbb{X} \rightrightarrows \mathbb{Y}$, which is defined as

$$\text{epi}_f(x) = f(x) + \mathbb{Y}_+$$

or, equivalently, by the condition

$$\text{grph}(\text{epi}_f) = \text{epi}(f),$$

where $\text{epi}(f) = \{(x,y) \in \mathbb{X} \times \mathbb{Y} : \ f(x) \le_\mathbb{Y} y\}$, is set-covering in the sense of Definition 2.1, with constant α/γ. To see this, fix $x \in \mathbb{X}$ and $r > 0$, and consider an element $\check{y} \in \mathbb{Y}$ as in (2.4), namely such that

$$\check{y} \le_\mathbb{Y} y, \quad \forall y \in B(f(x), \alpha r) \quad \text{and} \quad \|\check{y} - f(x)\| = \gamma\alpha r.$$

This implies that

$$B(f(x), \alpha r) \subseteq \check{y} + \mathbb{Y}_+.$$

Since it is $\check{y} \in B(f(x), \alpha\gamma r)$, then by virtue of (2.5) there must exist $u \in B(x, \gamma r)$ such that $\check{y} = f(u)$. Thus, letting $\rho = \gamma r$, one obtains

$$\begin{aligned}
B\left(\text{epi}_f(x), \frac{\alpha}{\gamma}\rho\right) &= B(f(x) + \mathbb{Y}_+, \alpha r) \subseteq B(f(x), \alpha r) + \mathbb{Y}_+ \\
&\subseteq \check{y} + \mathbb{Y}_+ = f(u) + \mathbb{Y}_+ = \text{epi}_f(u),
\end{aligned}$$

with $u \in \mathrm{B}(x, \rho)$. Whenever $(\mathbb{X}, \| \cdot \|)$ and $(\mathbb{Y}, \| \cdot \|)$ are, in particular, Banach spaces and $f : \mathbb{X} \longrightarrow \mathbb{Y}$ is a bounded linear operator, then according to the Banach-Schauder open mapping theorem f is known to be covering on $\mathbb{X}$ iff it is onto. In such an event, the quantity

$$\|f^{-1}\|_{-} = \sup_{y \in \mathbb{B}} \inf \{\|x\| : \ f(x) = y\} = \sup_{y \in \mathbb{B}} \mathrm{dist}\left(\mathbf{0}, f^{-1}(y)\right),$$

where $f^{-1} : \mathbb{Y} \rightrightarrows \mathbb{X}$ is the (generally) set-valued inverse mapping of f, is a positive element of $\mathbb{R}$ and, as a covering constant, one can take $\alpha = 1/\|f^{-1}\|_{-}$. Thus the epigraphical mapping of a bounded linear operator, which is onto, is set-covering with constant $1/\gamma \|f^{-1}\|_{-}$.

To the aim of providing further examples of classes of set-covering mappings, let us focus now on convex processes. The idea of a convex process is due to R.T. Rockafellar (see [31]) and emerges when dealing with derivatives of set-valued mappings or with certain constraint systems arising in optimization problems. After him, a set-valued mapping $\Theta : \mathbb{X} \rightrightarrows \mathbb{Y}$ between normed spaces is said to be a convex process if $\mathrm{grph}(\Theta)$ is a convex cone of $\mathbb{X} \times \mathbb{Y}$ with apex at the null vector, or, equivalently, iff Θ satisfies all the following three requirements:

(i) $\mathbf{0} \in \Theta(\mathbf{0})$;
(ii) $\Theta(\lambda x) = \lambda \Theta(x), \quad \forall \lambda > 0, \ \forall x \in \mathbb{X}$;
(iii) $\Theta(x_1) + \Theta(x_2) \subseteq \Theta(x_1 + x_2), \quad \forall x_1, x_2 \in \mathbb{X}$.

Clearly Θ is a convex process iff Θ^{-1} is so. Further, a convex process is said to be closed provided that so is its graph. A way to approach the study of the covering behaviour of convex processes is through the notion of openness at $\mathbf{0}$. According to [29], a convex process Θ is said to be *open* at $\mathbf{0}$ if there exists $\alpha > 0$ such that

$$(2.6) \qquad\qquad\qquad \Theta(\mathrm{int}\,\mathbb{B}) \supseteq \mathrm{int}\,\alpha\mathbb{B}.$$

Such a condition has been characterized in terms of finiteness of the inner norm of the inverse mapping. More precisely, defined the inner norm of a mapping $\Theta : \mathbb{X} \rightrightarrows \mathbb{Y}$ as

$$\|\Theta\|_{-} = \sup_{x \in \mathrm{dom}\,\Theta \cap \mathbb{B}} \inf \{\|y\| : \ y \in \Theta(x)\},$$

it has been established that Θ is open at $\mathbf{0}$ iff $\|\Theta^{-1}\|_{-} < +\infty$ and, as a constant appearing in (2.6), it is possible to take any value $\alpha \in (0, 1/\|\Theta^{-1}\|_{-}^{-1})$ (see [29]). Whenever $(\mathbb{X}, \| \cdot \|)$ and $(\mathbb{Y}, \| \cdot \|)$ are, in particular, Banach spaces, a sufficient condition for a closed, convex process $\Theta : \mathbb{X} \rightrightarrows \mathbb{Y}$ to be open at $\mathbf{0}$ is that Θ is onto (open mapping theorem for convex processes, see [3, 12, 29]).

The next result and the subsequent related remark show that a proper subclass of onto and closed convex processes can be found, whose elements are set-covering mappings.

Proposition 2.7. *Let $\Theta : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a closed, convex process between Banach spaces. If the following condition holds*

$$(2.7) \qquad\qquad \exists \alpha > 0, \quad \exists u \in \mathbb{B} \quad such \ that \quad \Theta(u) \supseteq \mathrm{int}\,\alpha\mathbb{B},$$

then Θ is set-covering with any constant $\tilde{\alpha} \in (0, \alpha)$. Vice versa, if Θ is set-covering with a constant $\alpha > 0$, then condition (2.7) holds.

Proof. Assume first that condition (2.7) holds true. Fix arbitrary $x \in \mathbb{X}$, $r > 0$ and $\tilde{\alpha} \in (0, \alpha)$, take $y \in \mathrm{B}(\Theta(x), \tilde{\alpha}r)$ and pick $\epsilon > 0$ in such a way that $\tilde{\alpha}(1 + \epsilon) < \alpha$. This implies that there exists of $v \in \Theta(x)$ such that

$$\|y - v\| < \tilde{\alpha}(1 + \epsilon)r,$$

that is $y - v \in \mathrm{int}\,\tilde{\alpha}(1+\epsilon)r\mathbb{B}$. Since, as a convex process, Θ is positively homogeneous, condition (2.7) entails the existence of $u \in r\mathbb{B}$ such that

$$\Theta(u) \supseteq \mathrm{int}\,\alpha r\mathbb{B} \supseteq \mathrm{int}\,\tilde{\alpha}(1 + \epsilon)r\mathbb{B}.$$

Consequently, one has

$$y = (y - v) + v \in \mathrm{int}\,\alpha r\mathbb{B} + \Theta(x) \subseteq \Theta(u) + \Theta(x) \subseteq \Theta(u + x),$$

where $u + x \in \mathrm{B}(x, r)$. It should be noticed that, according to the inclusion (2.7), the element u in $r\mathbb{B}$ does not depend neither on y nor on v.

Vice versa, if choosing $x = \mathbf{0}$ and $r = 1$ in Definition 2.1, since it is $\mathbf{0} \in \Theta(\mathbf{0})$, one finds

$$\mathrm{int}\,\alpha\mathbb{B} \subseteq \mathrm{B}(\Theta(\mathbf{0}), \alpha) \subseteq \Theta(u)$$

for some $u \in \mathbb{B}$ and hence condition (2.7) is verified at once. $\qquad\square$

Remark 2.8. Condition (2.7) is a quantitative requirement about the surjective behaviour of Θ that can be expressed in merely topological terms as

(2.8) $\mathrm{int}\,\Theta(\mathbf{0}) \neq \varnothing.$

Indeed, such a condition obviously implies (2.7). Vice versa, if u and α are as in (2.7), one has

$$\Theta(\mathbf{0}) = \Theta(u) + \Theta(-u) \supseteq \mathrm{int}\,\alpha\mathbb{B} + \Theta(-u),$$

wherefrom the interior nonemptiness condition in (2.8) follows. It is worth noting that condition (2.7) is essentially stronger than openness at $\mathbf{0}$. In other words, the latter is sufficient for a closed, convex process to be covering, whereas it does not in the case of the set-covering property, even in the case of convex processes. This occurence is illustrated in the next example.

Example 2.9. Consider the Banach space $\mathbb{X} = \mathbb{Y} = (\ell_p(\mathbb{N}), \|\cdot\|_p)$, with $p \in [1, +\infty)$ and define $\Theta : \ell_p(\mathbb{N}) \rightrightarrows \ell_p(\mathbb{N})$ as

$$\Theta(x) = x + \mathbb{X}_+,$$

where $\mathbb{X}_+ = \{x = (\xi_n)_{n \in \mathbb{N}} : \xi_n \geq 0, \quad \forall n \in \mathbb{N}\}$. It is not difficult to check that Θ is a closed, convex process and it is clear that Θ is also onto. Therefore Θ is open at $\mathbf{0}$ and hence covering. Nevertheless, in the light of Proposition 2.7 Θ fails to be set-covering. Indeed, it is $\Theta(\mathbf{0}) = \mathbb{X}_+$ and such a cone is well known to have empty topological interior (see, for instance, [20]).

2.2. **Related properties.** In the current subsection, the context is again that of set-valued mappings between metric spaces. Given a set-valued mapping $\Psi : X \rightrightarrows Y$ and a nonempty set $S \subseteq Y$, there are two natural notions of inverse image of S through Ψ, which are known in set-valued analysis as upper (or strong) and lower (or weak) inverse, respectively. Here yet another notion is considered, which is defined as follows

$$\Psi^{-\sharp}(S) = \{x \in X : \ S \subseteq \Psi(x)\}.$$

Letting S to vary in 2^Y, one obtains from Ψ a set-valued mapping $\Psi^{-\sharp} : 2^Y \rightrightarrows X$. Such a mapping can be viewed as solution mapping of a set-inclusion, where the set S plays the role of a parameter. As one expects, a reformulation of the set-covering property can be expressed in terms of error bound for the solution mapping associated with such a parameterized set-inclusion problem.

Proposition 2.10. *Let $\Psi : X \rightrightarrows Y$ be a set-valued mapping between metric spaces.*

(i) *If Ψ is set-covering with constant $\alpha > 0$, then*

$$(2.9) \qquad \operatorname{dist}\left(x, \Psi^{-\sharp}(S)\right) \leq \frac{1}{\alpha}\operatorname{exc}\left(S, \Psi(x)\right), \quad \forall x \in X, \ \forall S \in 2^Y.$$

(ii) *If inequality (2.9) holds, then Ψ is set-covering with any constant $\tilde{\alpha} \in (0, \alpha)$.*

Proof. (i) Assume that $x \in X$ and $S \subseteq Y$ are arbitrary. If $\operatorname{exc}\left(S, \Psi(x)\right) = +\infty$, then (2.9) trivially holds. In this case, it may happen that $\Psi^{-\sharp}(S) = \varnothing$. Thus, let us pass to consider the case $r = \operatorname{exc}\left(S, \Psi(x)\right) < +\infty$. Since Ψ takes closed values, if $r = 0$ one has $S \subseteq \Psi(x)$ and hence $x \in \Psi^{-\sharp}(S)$. If $r > 0$, since $S \subseteq \mathrm{B}(\Psi(x), r)$, according to Definition 2.1, there exists $u \in \mathrm{B}(x, r/\alpha)$ such that $S \subseteq \Psi(u)$. It follows

$$\operatorname{dist}\left(x, \Psi^{-\sharp}(S)\right) \leq d(x, u) \leq \frac{1}{\alpha}\operatorname{exc}\left(S, \Psi(x)\right).$$

(ii) Fix any $\tilde{\alpha} \in (0, \alpha)$. To see that in this case Definition 2.1 is satisfied, it suffices to take $S = \mathrm{B}(\Psi(x), \tilde{\alpha}r)$ and to observe that, with this choice, it is $\operatorname{exc}\left(S, \Psi(x)\right) \leq \tilde{\alpha}r < +\infty$. Therefore, from inequality (2.9) one gets

$$\operatorname{dist}\left(x, \Psi^{-\sharp}(S)\right) \leq \frac{1}{\alpha}\tilde{\alpha}r < r,$$

which means that there exists $u \in \mathrm{B}(x, r)$ such that $\Psi(u) \supseteq \mathrm{B}(\Psi(x), \tilde{\alpha}r)$. $\qquad \square$

Remark 2.11. It is proper to warn the reader that, in general, it may happen that $\operatorname{dom}\Psi^{-\sharp} \neq 2^Y$. From inequality (2.9) it follows that, whenever Ψ is set-covering and $S \subseteq Y$ is such that $\operatorname{exc}\left(S, \Psi(x)\right) < +\infty$, then it must be $\Psi^{-\sharp}(S) \neq \varnothing$. In particular, $\operatorname{dom}\Psi^{-\sharp}$ includes all bounded subsets of Y. Let us denote by $\mathcal{B}(Y)$ the collection of all such subsets of Y.

The next step consists in linking the set-covering property with the Lipschitzian behaviour of the mapping $\Psi^{-\sharp}$. To this aim, the set $\mathcal{B}(Y)$ is equipped with the Hausdorff distance.

Proposition 2.12. *Let $\Psi : X \rightrightarrows Y$ be a set-valued mapping between metric spaces. If Ψ is set-covering with constant $\alpha > 0$, then $\Psi^{-\sharp} : \mathcal{B}(Y) \rightrightarrows X$ is Lipschitz with constant $1/\alpha$. Vice versa, given $\Psi : X \longrightarrow \mathcal{B}(Y)$, if $\Psi^{-\sharp} : \mathcal{B}(Y) \rightrightarrows X$ is Lipschitz*

with constant $0 < l < +\infty$, *then* Ψ *is set-covering on* X *with any constant* $\alpha \in (0, 1/l)$.

Proof. Consider a pair of elements $A, B \in \mathcal{B}(Y)$. From assertion (i) in Proposition 2.10, one has

$$\mathrm{dist}\left(x, \Psi^{-\sharp}(B)\right) \leq \alpha^{-1}\mathrm{exc}\left(B, \Psi(x)\right), \quad \forall x \in X.$$

Thus, for all those $x \in \Psi^{-\sharp}(A)$, i.e. such that $A \subseteq \Psi(x)$, one finds

$$\mathrm{dist}\left(x, \Psi^{-\sharp}(B)\right) \leq \alpha^{-1}\mathrm{exc}\left(B, A\right),$$

whence

$$\mathrm{exc}\left(\Psi^{-\sharp}(A), \Psi^{-\sharp}(B)\right) = \sup_{x \in \Psi^{-\sharp}(A)} \mathrm{dist}\left(x, \Psi^{-\sharp}(B)\right) \leq \alpha^{-1}\mathrm{exc}\left(B, A\right).$$

To achieve the inequality

$$\mathrm{Haus}(\Psi^{-\sharp}(A), \Psi^{-\sharp}(B)) \leq \alpha^{-1}\mathrm{Haus}(A, B),$$

it suffices to interchange the role of A and B.

To prove the second assertion in the thesis, let $x \in X$ and $r > 0$ be arbitrary. Since now $\Psi(x) \in \mathcal{B}(Y)$, the same is true for $\mathrm{B}(\Psi(x), \alpha r)$. By the Lipschitz continuity of $\Psi^{-\sharp}$ with constant l, if taking any $\alpha \in (0, 1/l)$ one finds

$$\begin{aligned}\mathrm{Haus}(\Psi^{-\sharp}(\Psi(x)), \Psi^{-\sharp}(\mathrm{B}(\Psi(x), \alpha r))) &\leq l\mathrm{Haus}(\Psi(x), \mathrm{B}(\Psi(x), \alpha r)) \\ &= l\mathrm{exc}\left(\mathrm{B}(\Psi(x), \alpha r), \Psi(x)\right) < r.\end{aligned}$$

On the other hand, observe that it is

$$\begin{aligned}\Psi^{-\sharp}(\Psi(x)) &= \{u \in X : \ \Psi(x) \subseteq \Psi(u)\} \supseteq \Psi^{-\sharp}(\mathrm{B}(\Psi(x), \alpha r)) \\ &= \{u \in X : \ \mathrm{B}(\Psi(x), \alpha r) \subseteq \Phi(u)\}.\end{aligned}$$

Thus, the last inequality amounts to say that

$$\mathrm{exc}\left(\Psi^{-\sharp}(\Psi(x)), \Psi^{-\sharp}(\mathrm{B}(\Psi(x), \alpha r))\right) < r,$$

wherefrom, as in particular it is $x \in \Psi^{-\sharp}(\Psi(x))$, one obtains

$$\mathrm{dist}\left(x, \{u \in X : \ \mathrm{B}(\Psi(x), \alpha r) \subseteq \Phi(u)\}\right) < r.$$

The last inequality implies the existence of $u \in \mathrm{B}(x, r)$ with the property that $\Psi(u) \supseteq \mathrm{B}(\Psi(x), \alpha r)$, thereby showing that Ψ is set-covering with constant α. $\square$

2.3. **Stability of set-covering.** The present subsection is devoted to illustrating some robustness features of the set-covering property in the presence of various kinds of perturbation.

Proposition 2.13. *Let* $\Psi : X \rightrightarrows Y$ *be set-covering on* X *with constant* α *and let* $g : Y \longrightarrow Z$ *be covering on* Y *with constant* β. *Then, their composition* $g \circ \Psi : X \rightrightarrows Z$ *is set-covering with any constant* $\gamma \in (0, \alpha\beta)$.

Proof. By the set-covering property of Ψ, corresponding to $x \in X$ and $r > 0$, there exists $u \in \mathrm{B}(x, r)$ such that $\Psi(u) \supseteq \mathrm{B}(\Psi(x), \alpha r)$, whence it follows

$$(2.10) \qquad (g \circ \Psi)(u) \supseteq g(\mathrm{B}(\Psi(x), \alpha r)).$$

By the covering property on Y of g, one has

$$(2.11) \qquad g(\mathrm{B}(y, \alpha r)) \supseteq \mathrm{B}(g(y), \alpha \beta r), \quad \forall y \in Y, \ \forall r > 0.$$

Notice that, since for any $z \in Z$ it is

$$\mathrm{dist}\,(z, (g \circ \Psi)(x)) = \inf_{y \in \Psi(x)} d(z, g(y)),$$

it holds

$$\mathrm{B}((g \circ \Psi)(x), \gamma r) \subseteq \bigcup_{y \in \Psi(x)} \mathrm{B}(g(y), \alpha \beta r).$$

Thus, in the light of inclusion (2.11), one obtains

$$\begin{aligned} \mathrm{B}((g \circ \Psi)(x), \gamma r) \ &\subseteq\ \bigcup_{y \in \Psi(x)} g(\mathrm{B}(y, \alpha r)) \subseteq g \left(\bigcup_{y \in \Psi(x)} \mathrm{B}(y, \alpha r) \right) \\ &\subseteq\ g(\mathrm{B}(\Psi(x), \alpha r)). \end{aligned}$$

By recalling inclusion (2.10), one deduces that

$$(g \circ \Psi)(u) \supseteq \mathrm{B}((g \circ \Psi)(x), \gamma r).$$

This completes the proof. $\qquad\square$

The next proposition relates to a stability phenomenon regarding set-covering, which can be observed to take place in the presence of additive perturbations by multivalued Lipschitz mappings. In doing so, along with the previous one, it provides as well a tool for building further examples of classes of set-covering mappings.

Proposition 2.14. *Let X be a metric space and let Y be a vector space, equipped with a shift invariant metric. Suppose that $\Psi : X \rightrightarrows Y$ is set-covering on X with constant $\alpha > 0$, whereas $\Phi : X \rightrightarrows Y$ is Lipschitz on X with constant $\beta \in [0, \alpha)$. Then, for every $\epsilon \in (0, \alpha - \beta)$, the mapping $\Psi + \Phi$ is set-covering on X with constant $\alpha - \beta - \epsilon$.*

Proof. Fixed $\epsilon \in (0, \alpha - \beta)$, $x \in X$ and $r > 0$, according to Definition 2.1 one has to show that there exists $x' \in \mathrm{B}(x, r)$ such that

$$(2.12) \qquad \Psi(x') + \Phi(x') \supseteq \mathrm{B}(\Psi(x) + \Phi(x), (\alpha - \beta - \epsilon)r).$$

Since Ψ is covering on X with constant α, there exists $x' \in \mathrm{B}(x, r)$ such that $\Psi(x') \supseteq \mathrm{B}(\Psi(x), \alpha r)$, and hence

$$(2.13) \qquad \Psi(x') + \Phi(x') \supseteq \mathrm{B}(\Psi(x), \alpha r) + \Phi(x').$$

Since Φ is Lipschitz on X with constant $\beta \in [0, \alpha)$, one has

$$\mathrm{exc}\,\big(\Phi(x), \Phi(x')\big) \leq \beta d(x, x') \leq \beta r.$$

Furthermore, as the metric defined on Y is shift invariant, it holds

$$
\begin{aligned}
\text{exc}\left(\Psi(x) + \Phi(x), \Psi(x) + \Phi(x')\right) &= \sup_{\substack{u_1 \in \Psi(x) \\ w_1 \in \Phi(x)}} \text{dist}\left(u_1 + w_1, \Psi(x) + \Phi(x')\right) \\
&= \sup_{\substack{u_1 \in \Psi(x) \\ w_1 \in \Phi(x)}} \inf_{\substack{u_2 \in \Psi(x) \\ w_2 \in \Phi(x')}} d(u_1 + w_1, u_2 + w_2) \\
&\leq \sup_{\substack{u_1 \in \Psi(x) \\ w_1 \in \Phi(x)}} \inf_{w_2 \in \Phi(x')} d(u_1 + w_1, u_1 + w_2) \\
&= \sup_{w_1 \in \Phi(x)} \inf_{w_2 \in \Phi(x')} d(w_1, w_2) \\
&= \text{exc}\left(\Phi(x), \Phi(x')\right).
\end{aligned}
$$

Consequently, one has

$$(2.14) \qquad \text{B}(\Psi(x) + \Phi(x'), (\alpha - \epsilon)r) \supseteq \text{B}(\Psi(x) + \Phi(x), (\alpha - \beta - \epsilon)r).$$

Indeed, from the last inequalities one gets

$$
\begin{aligned}
\text{dist}\left(v, \Psi(x) + \Phi(x')\right) &\leq \text{dist}\left(v, \Psi(x) + \Phi(x)\right) \\
&+ \text{exc}\left(\Psi(x) + \Phi(x), \Psi(x) + \Phi(x')\right) \\
&\leq (\alpha - \beta - \epsilon)r + \beta r = (\alpha - \epsilon)r.
\end{aligned}
$$

Let us show now that

$$(2.15) \qquad \text{B}(\Psi(x), \alpha r) + \Phi(x') \supseteq \text{B}(\Psi(x) + \Phi(x'), (\alpha - \epsilon)r).$$

Taken an arbitrary $v \in \text{B}(\Psi(x) + \Phi(x'), (\alpha - \epsilon)r)$, there exist $u_\epsilon \in \Psi(x)$ and $w_\epsilon \in \Phi(x')$ such that $d(v, u_\epsilon + w_\epsilon) < \alpha r$. Thus, it results in

$$v = v - w_\epsilon + w_\epsilon,$$

where, again by virtue of the shift invariance of the metric on Y, it is

$$\text{dist}\left(v - w_\epsilon, \Psi(x)\right) = \text{dist}\left(v, \Psi(x) + w_\epsilon\right) \leq d(v, u_\epsilon + w_\epsilon) < \alpha r,$$

so $v \in \text{B}(\Psi(x), \alpha r) + \Phi(x')$. In order to achieve the required condition (2.12), it suffices to combine inclusions (2.13), (2.14), and (2.15). This completes the proof. $\qquad\square$

Remark 2.15. (i) The phenomenon described by Proposition 2.14 can be inserted in the framework of the stability analysis for covering behaviours started with the well-known Milyutin theorem (see, for instance, [1, 2, 9, 10]). Note that, in contrast to the latter, which refers to a traditional covering behaviour, no completeness assumption is needed in the case of set-covering, because no iteration process is needed. In fact, in the case of standard covering, the unperturbed mapping covers the required ball by the image of a *whole ball*, while the perturbation shifts *separately each point* of this image. As a consequence, the perturbed image may have "holes" so, to cover the required ball, one needs to use an iteration process. In the case of set-covering, while the image of a *single point* under the unperturbed mapping covers the required ball, the perturbation just slightly shifts this image *as the whole*, so it still covers a ball with a properly reduced radius.

(ii) As a comment to the assumptions of Proposition 2.14, it is to be pointed out that the shift invariance requirement on the metric of Y is not really restrictive. Indeed, according to a result due to S. Kakutani, any linear metric space can be equivalently remetrized by a shift invariant distance (see [33], Theorem 2.2.11).

Example 2.16. Let (X, d) be a metric space and let (Y, d) be a vector space endowed with a shift-invariant metric. Then, as a consequence of Proposition 2.14 and Example 2.3, if $\delta : X \longrightarrow [0, +\infty)$ fulfils condition (2.3) and $g : X \longrightarrow Y$ is a Lipschitz mapping with constant $\beta < \alpha_\delta$, then the set-valued mapping $\Psi : X \rightrightarrows Y$ given by

$$\Psi(x) = \mathrm{B}(g(x), \delta(x)) = \mathrm{B}(\mathbf{0}, \delta(x)) + g(x)$$

turns out to be set-covering with any constant $\alpha \in (0, \alpha_\delta - \beta)$.

3. SET-INCLUSION POINTS OF PAIRS OF MAPPINGS AND APPLICATIONS

3.1. **Set-inclusion points.** The next definition introduces a very general problem, which can be posed whenever any pair of multivalued mappings is given.

Definition 3.1. Given two set-valued mappings $\Psi : X \rightrightarrows Y$ and $\Phi : X \rightrightarrows Y$, an element $x \in X$ is called a *set-inclusion point* of the (ordered) pair (Φ, Ψ) if

$$\Phi(x) \subseteq \Psi(x).$$

Denote

$$\mathrm{Inc}(\Phi, \Psi) = \{x \in X : \ \Phi(x) \subseteq \Psi(x)\}.$$

A set-inclusion point problem is an abstract formalism able to subsume in its extreme generality several specific problems, having or not a variational nature. For instance, it enables one to embed equilibrium conditions, fixed or coincidence point problems, generalized equations, by proper choices of Φ and Ψ. Nonetheless, it is when Φ (and hence Ψ) is actually a multivalued mapping that its peculiarity appears. Besides, it is interesting to note that $\mathrm{Inc}(\Phi, \Psi)$ can be regarded as the set of all fixed points of the mapping $\Psi^{-\sharp} \circ \Phi : X \rightrightarrows X$. A question related to the set-inclusion problem was considered in [16], where conditions for the existence of a set $K \subseteq X$ such that $\Psi(K) = \Phi(K)$ (called coincidence set) were established.

Remark 3.2. It is worth noting that, as a straightforward consequence of Lemma 1.2, the set $\mathrm{Inc}(\Phi, \Psi)$ is closed whenever Ψ is Hausdorff u.s.c. on X and Φ is Lipschitz on X. Indeed, if $\mathrm{Inc}(\Phi, \Psi) \neq \varnothing$ and $(x_n)_{n \in \mathbb{N}}$ is a sequence in $\mathrm{Inc}(\Phi, \Psi)$ converging to x_0 as $n \to \infty$, one finds

$$0 = \liminf_{n \to \infty} \mathrm{exc}_{\Phi, \Psi}(x_n) \geq \mathrm{exc}_{\Phi, \Psi}(x_0) \geq 0,$$

wherefrom, by closedness of $\Psi(x_0)$, it follows that $x_0 \in \mathrm{Inc}(\Phi, \Psi)$.

In what follows, pursuing a similar line of research as in [1, 2], the question of the solution existence of set-inclusion points is analyzed in the general setting of multifunctions between metric spaces. In particular, the next result provides a sufficient condition for a set-inclusion problem, involving a set-covering and a Lipschitz mappings, to admit a solution, as well as an error bound for its solution set.

Theorem 3.3. *Let* $\Psi : X \rightrightarrows Y$ *and* $\Phi : X \longrightarrow \mathcal{B}(Y)$ *be given set-valued mappings between metric spaces. Suppose that:*

(i) (X, d) *is metrically complete;*

(ii) Ψ *is Hausdorff u.s.c. and set-covering on* X, *with constant* $\alpha > 0$;

(iii) Φ *is Lipschitz on* X *with constant* $\beta \in [0, \alpha)$.

Then, $\mathrm{Inc}(\Phi, \Psi) \neq \varnothing$ *and the following estimate holds*

$$(3.1) \qquad \mathrm{dist}\,(x, \mathrm{Inc}(\Phi, \Psi)) \leq \frac{\mathrm{exc}\,(\Phi(x), \Psi(x))}{\alpha - \beta}, \quad \forall x \in\ X.$$

Proof. The proof is based on a variational technique. Notice indeed that, in order to prove the existence of a set-inclusion point $\bar{x} \in X$ for the pair Ψ and Φ, it suffices to show that the function $\mathrm{exc}_{\Phi,\Psi} : X \longrightarrow [0, +\infty)$ attains the value 0 at some point $\bar{x}$. This, because the validity of $\mathrm{dist}\,(y, \Psi(\bar{x})) = 0$ for every $y \in \Phi(\bar{x})$, as $\Psi(\bar{x})$ is a closed set, implies $\Phi(\bar{x}) \subseteq \Psi(\bar{x})$. The nonemptiness of the values taken by Ψ, along with the boundedness of the values of Φ, make the function $\mathrm{exc}_{\Phi,\Psi}$ real valued all over X. So, take an arbitrary $x_0 \in X$. In the case $\mathrm{exc}_{\Phi,\Psi}(x_0) = 0$, one has immediately $x_0 \in \mathrm{Inc}(\Phi, \Psi) \neq \varnothing$ and the estimate in (3.1). So, assume henceforth that $\mathrm{exc}_{\Phi,\Psi}(x_0) > 0$. Observe that, under the current hypotheses, the function $\mathrm{exc}_{\Phi,\Psi}$ turns out to be l.s.c. on X, by virtue of Lemma 1.2. As it is obviously bounded from below and X is complete, then the Ekeland variational principle applies. Accordingly, for every $\lambda > 0$ there exists $x_\lambda \in X$ such that

$$\mathrm{exc}_{\Phi,\Psi}(x_\lambda) \leq \mathrm{exc}_{\Phi,\Psi}(x_0),$$

$$(3.2) \qquad d(x_\lambda, x_0) \leq \lambda,$$

$$(3.3) \qquad \mathrm{exc}_{\Phi,\Psi}(x_\lambda) < \mathrm{exc}_{\Phi,\Psi}(x) + \frac{\mathrm{exc}_{\Phi,\Psi}(x_0)}{\lambda} d(x, x_\lambda), \quad \forall x \in X \backslash \{x_\lambda\}.$$

Take $\lambda = \mathrm{exc}_{\Phi,\Psi}(x_0)/(\alpha - \beta)$. The claim to be proved is that $\mathrm{exc}_{\Phi,\Psi}(x_\lambda) = 0$. Ab absurdo, assume that $r_\lambda = \mathrm{exc}_{\Phi,\Psi}(x_\lambda) > 0$. Since $\Phi(x_\lambda) \subseteq \mathrm{B}(\Psi(x_\lambda), r_\lambda)$, the set-covering property of Ψ enables one to state the existence of $u \in \mathrm{B}(x_\lambda, r_\lambda/\alpha)$ such that

$$\Psi(u) \supseteq \mathrm{B}(\Psi(x_\lambda), r_\lambda) \supseteq \Phi(x_\lambda).$$

Therefore, recalling that the function $x \mapsto \mathrm{exc}\,(\Phi(x), \Psi(u))$ is Lipschitz on X, with constant β, as a consequence of the assumption on Φ and Remark 1.1, one obtains

$$(3.4) \qquad \mathrm{exc}_{\Phi,\Psi}(u) \leq \mathrm{exc}\,(\Phi(x_\lambda), \Psi(u)) + \beta d(u, x_\lambda) = \beta d(u, x_\lambda).$$

Notice that it must be $u \in X \backslash \{x_\lambda\}$, otherwise it would be $\mathrm{exc}_{\Phi,\Psi}(x_\lambda) = 0$. By choosing $x = u$ in inequality (3.3) and taking into account inequality (3.4), one finds

$$\mathrm{exc}_{\Phi,\Psi}(x_\lambda) < \mathrm{exc}_{\Phi,\Psi}(u) + (\alpha - \beta)d(u, x_\lambda) \leq \alpha d(u, x_\lambda) \leq r_\lambda,$$

which leads to an evident contradiction. This allows one to conclude that $\mathrm{exc}_{\Phi,\Psi}(x_\lambda) = 0$, and hence $x_\lambda \in \mathrm{Inc}(\Phi, \Psi) \neq \varnothing$. From inequality (3.2), it readily follows

$$\mathrm{dist}\,(x_0, \mathrm{Inc}(\Phi, \Psi)) \leq d(x_0, x_\lambda) \leq \frac{\mathrm{exc}\,(\Phi(x_0), \Psi(x_0))}{\alpha - \beta}.$$

By arbitrariness of x_0, this completes the proof. $\qquad \square$

When Φ is a single-valued mapping, Theorem 3.3 allows one to obtain, as a special case, a well-known result about the existence and error bound estimates for coincidence points of the inclusion $\Phi(x) \in \Psi(x)$. Nevertheless, in order to achieve such a result, the conventional notion of covering is actually enough (see [1]). In contrast to this, as far as set-inclusion points are concerned, the set-covering property plays an essential role. The following counterexample shows that such a property can not be replaced with the usual covering notion for set-valued mappings.

Example 3.4. Consider the set-valued mapping $\Psi : \mathbb{R} \rightrightarrows \mathbb{R}^2$, introduced in Example 2.2, that is covering with constant $\alpha = 1$. It is easy to check that this mapping is also Hausdorff u.s.c. on $\mathbb{R}$. Define a further mapping $\Phi : \mathbb{R} \rightrightarrows \mathbb{R}^2$ as follows

$$\Phi(x) = \left(\frac{|x|}{2} + 1 \right) \mathbb{B}.$$

Since it is

$$\mathrm{exc}\,(\Phi(x_1), \Phi(x_2)) \leq \frac{1}{2}|x_1 - x_2|, \quad \forall x_1,\, x_2 \in \mathbb{R},$$

Φ turns out to be Lipschitz on $\mathbb{R}$ with constant $\beta = 1/2 < 1 = \alpha$ and bounded value. Nonetheless, in this case it happens that $\mathrm{Inc}(\Phi, \Psi) = \varnothing$, as one readily observes, being $\mathrm{int}\,\Psi(x) = \varnothing$ for every $x \in \mathbb{R}$.

A related application of the notion of set-covering concerns the fixed point theory for multivalued mappings.

Definition 3.5. An element x of a metric space X is said to be a *strongly fixed point* of a set-valued mapping $\Psi : X \rightrightarrows X$ if for some $r > 0$ it is

$$\mathrm{B}(x, r) \subseteq \Psi(x).$$

The set of all strongly fixed point of Ψ is denoted henceforth by $\mathrm{SFix}(\Psi)$.

In the following proposition, strongly fixed points are shown to arise in connection with set-covering mappings with constant greater than 1 (a sort of expanding mappings).

Proposition 3.6. *Let $\Psi : X \rightrightarrows X$ be a set-valued mapping defined on a vector space, endowed with a complete and shift invariant metric. If Ψ is u.s.c. and set-covering on X, with constant $\alpha > 1$, then $\mathrm{SFix}(\Psi) \neq \varnothing$ and it holds*

$$\mathrm{dist}\,(x, \mathrm{SFix}(\Psi)) \leq \frac{\mathrm{dist}\,(x, \Psi(x))}{\alpha - 1}, \quad \forall x \in X.$$

Moreover, $\mathrm{SFix}(\Psi)$ is a dense subset of the set of all fixed points of Ψ.

Proof. Fix arbitrary $x_0 \in X$ and $r > 0$ and consider the set-valued mapping $\Phi_r : X \rightrightarrows X$ given by

$$\Phi_r(x) = \mathrm{B}(x, r).$$

Let us show that, under the current hypotheses, Φ_r is Lipschitz with constant $\beta = 1$. Notice indeed that, by the shift invariance of the metric on X, one has $\mathrm{B}(x, r) = x + \mathrm{B}(\mathbf{0}, r)$. Thus, taken $x_1, x_2 \in X$, if $y_1 \in \Phi_r(x_1) = x_1 + \mathrm{B}(\mathbf{0}, r)$, for some $u \in \mathrm{B}(\mathbf{0}, r)$ it results in

$$\mathrm{dist}\,(y_1, \Phi_r(x_2)) \quad = \quad \mathrm{dist}\,(y_1, x_2 + \mathrm{B}(\mathbf{0}, r)) = \mathrm{dist}\,(y_1 - x_2, \mathrm{B}(\mathbf{0}, r))$$

$$
\begin{aligned}
&= \quad \mathrm{dist}\,(x_1 + u - x_2, \mathrm{B}(\mathbf{0}, r)) \\
&\leq \quad d(x_1, x_2) + \mathrm{dist}\,(u, \mathrm{B}(\mathbf{0}, r)) = d(x_1, x_2),
\end{aligned}
$$

whence

$$
\mathrm{exc}\,(\Phi_r(x_1), \Phi_r(x_2)) \leq d(x_1, x_2).
$$

Since $\alpha > 1$, then it is possible to apply Theorem 3.3, according to which $\mathrm{Inc}(\Phi_r, \Psi) \neq \varnothing$ and

$$
\mathrm{dist}\,(x_0, \mathrm{Inc}(\Phi_r, \Psi)) \leq \frac{\mathrm{exc}\,(\Phi_r(x_0), \Psi(x_0))}{\alpha - 1}.
$$

Now, observe that $\mathrm{SFix}(\Psi) = \cup_{r>0}\mathrm{Inc}(\Phi_r, \Psi) \neq \varnothing$. By using again the shift invariance of the metric, one obtains

$$
\begin{aligned}
\mathrm{dist}\,(x_0, \mathrm{SFix}(\Psi)) \quad &\leq \quad \inf_{r>0} \frac{\mathrm{exc}\,(\Phi_r(x_0), \Psi(x_0))}{\alpha - 1} \\
&= \quad \inf_{r>0} \sup_{u \in \mathrm{B}(\mathbf{0},r)} \frac{\mathrm{dist}\,(x_0 + u, \Psi(x_0))}{\alpha - 1} \\
&\leq \quad \inf_{r>0} \sup_{u \in \mathrm{B}(\mathbf{0},r)} \frac{d(x_0 + u, x_0) + \mathrm{dist}\,(x_0, \Psi(x_0))}{\alpha - 1} \\
&= \quad \frac{\mathrm{dist}\,(x_0, \Psi(x_0))}{\alpha - 1}.
\end{aligned}
$$

This proves the first assertion in thesis. The second one is a straightforward consequence of the first. $\qquad\square$

3.2. **Applications to exact penalization.** Of course, in force of its versatility, a set-inclusion problem associated with a given pair of multivalued mappings may appear among the constraints of optimization problems. In the remaining part of this section, it is shown how the error bound estimate provided in Theorem 3.3 can be exploited in deriving exact penalization results specific for problems with a constraint of this type. After [13, 36], exact penalization is a well-known approach for the treatment of variously constrained optimization problems, whose effectiveness is recognized from the theoretical as well as from the algorithmic viewpoint. Essentially, it consists in reducing a given constrained extremum problem to an unconstrained one, by replacing its objective functional with a so-called penalty functional, which is obtained by adding to the original objective functional a term properly quantifying the constraint violation. Relying on this idea, a comprehensive theory has been developed during the last decades. The reader is referred to [11] for a recent presentation, which unifies many results on exact penalization existing in the related literature (see also [37]). For convergence rate estimates for power penalty methods see [4]. Let us focus here on constrained optimization problems of the form

$$
(\mathcal{P}) \qquad\qquad \min \varphi(x) \quad \text{subject to} \quad x \in \mathcal{R} = \mathrm{Inc}(\Phi, \Psi),
$$

where the objective functional $\varphi : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ and the multivalued mappings $\Psi : X \rightrightarrows Y$ and $\Phi : X \longrightarrow \mathcal{B}(Y)$ are given problem data. The penalty functional

$$
\varphi_l(x) = \varphi(x) + l \cdot \mathrm{exc}_{\Phi, \Psi}(x)
$$

enables one to associate with problem $(\mathcal{P})$ the unconstrained problem

$$(\mathcal{P}_l) \qquad\qquad \min_{x \in X} \varphi_l(x).$$

Conditions ensuring that a local solution to $(\mathcal{P})$ is also a local solution to $(\mathcal{P}_l)$, provided that l is large enough, namely ensuring the existence of an exact penalty functional, turns out to be useful for formulating necessary optimality conditions for problem $(\mathcal{P})$. In the next result, a condition of this type is recalled. Results on the existence of exact penalty functions can be found also in [5], where sufficient conditions are derived in the absence of standard regularity conditions.

Theorem 3.7. *Let $\bar{x} \in \mathcal{R}$ be a local solution to $(\mathcal{P})$. Suppose that*

  (i) φ *is locally Lipschitz near $\bar{x}$, with constant $l_\varphi > 0$;*
  (ii) (X, d) *is metrically complete;*
  (iii) Ψ *is Hausdorff u.s.c. and set-covering on X, with constant $\alpha > 0$;*
  (iv) Φ *is Lipschitz on X with constant $\beta \in [0, \alpha)$.*

Then, the penalty functional φ_l is exact at $\bar{x}$ (i.e. $\bar{x}$ is an unconstrained local minimizer of φ_l), for every $l \geq \frac{l_\varphi}{\alpha - \beta}$.

Proof. The thesis can be readily achieved from Theorem 3.3, by applying the following well-known fact, often referred as to the exact penalization principle for Lipschitz functions: let $\varphi : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ be a function defined on a metric space, which is locally Lipschitz near $\bar{x}$ with constant $l_\varphi \geq 0$, and let $\gamma : X \longrightarrow [0, +\infty)$ be a function fulfilling for some $\kappa > 0$ the error bound

$$\text{dist}\left(x, \gamma^{-1}(0)\right) \leq \kappa \gamma(x),$$

for every x in a neighbourhood of $\bar{x}$. Then, if $\bar{x}$ is a local minimizer of φ subject to $\gamma(x) = 0$, it is also a unconstrained minimizer of the function

$$\varphi_s(x) = \varphi(x) + s\gamma(x),$$

provided that $s \geq l_\varphi \kappa$ (see, for instance, [14]). Indeed in the current case, as a residual function γ, it suffices to take $\gamma(x) = \text{exc}_{\Phi, \Psi}(x)$. $\qquad\square$

By exploiting again the above error bound estimate for set-inclusion points, it is possible to establish a sort of converse of the last result, which is valid for global solutions.

Proposition 3.8. *Let problem $(\mathcal{P})$ admit global solutions. Suppose that:*

  (i) φ *is Lipschitz on X with constant $l_\varphi > 0$;*
  (ii) (X, d) *is metrically complete;*
  (iii) Ψ *is Hausdorff u.s.c. and set-covering on X, with constant $\alpha > 0$;*
  (iv) Φ *is Lipschitz on X with constant $\beta \in [0, \alpha)$.*

Fix $\epsilon > 0$ and set $l_\epsilon = \frac{(1+\epsilon)l_\varphi}{\alpha - \beta}$. If $\hat{x}$ is a strict global solution to problem $(\mathcal{P}_{l_\epsilon})$, then $\hat{x}$ globally solves also $(\mathcal{P})$.

Proof. By virtue of the error bound estimate (3.1), corresponding to any $\epsilon > 0$ an element $x_\epsilon \in \mathcal{R}$ can be found such that

$$d(x_\epsilon, \hat{x}) \leq \frac{1+\epsilon}{\alpha - \beta} \text{exc}\left(\Phi(\hat{x}), \Psi(\hat{x})\right).$$

Denote by $\bar{x} \in \mathcal{R}$ a global solution of $(\mathcal{P})$. As $\hat{x}$ solves problem $(\mathcal{P}_{l_\epsilon})$, by the last inequality and the Lipschitz continuity of φ, one obtains

$$
\begin{aligned}
\varphi(\bar{x}) &= \varphi_{l_\epsilon}(\bar{x}) \geq \varphi_{l_\epsilon}(\hat{x}) = \varphi(\hat{x}) + l_\epsilon \mathrm{exc}\left(\Phi(\hat{x}), \Psi(\hat{x})\right) \\
&\geq \varphi(x_\epsilon) - l_\varphi d(\hat{x}, x_\epsilon) + l_\epsilon \mathrm{exc}\left(\Phi(\hat{x}), \Psi(\hat{x})\right) \geq \varphi(x_\epsilon) \geq \varphi(\bar{x}).
\end{aligned}
$$

The consequent fact that $\varphi_{l_\epsilon}(\hat{x}) = \varphi_{l_\epsilon}(x_\epsilon)$, since $\hat{x}$ is strict as a global solution to $(\mathcal{P}_{l_\epsilon})$, entails that $\hat{x} = x_\epsilon$, so that also $\hat{x} \in \mathcal{R}$. Thus, on account of the above inequalities, it is possible to conclude that $\hat{x}$ is a global solution of $(\mathcal{P})$. $\qquad\square$

Another approach to penalization methods in constrained optimization rests upon the concept of problem calmness, which was introduced by R.T. Rockafellar. This approach requires to regard a given problem as a particular specialization of a class of parameterized problems. In the case under study, the following class will be considered

$$(\mathcal{P}_p) \qquad \min \varphi(x) \quad \text{subject to} \quad x \in \mathcal{R}(p) = \{x \in X : \ \Phi(p, x) \subseteq \Psi(p, x)\},$$

where the data $\Phi : P \times X \longrightarrow \mathcal{B}(Y)$ and $\Psi : P \times X \rightrightarrows Y$ now depend also on $p \in P$, with (P, d) denoting a metric space of parameters. Notice that, unless suitable assumptions are introduced, one can not expect in general that $\mathrm{dom}\,\mathcal{R} = P$. With respect to this problem parameterization, the penalty functional $\varphi_l : P \times X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ becomes

$$\varphi_l(p, x) = \varphi(x) + l \cdot \mathrm{exc}\left(\Phi(p, x), \Psi(p, x)\right).$$

Definition 3.9. Let $\bar{p} \in P$ and let $\bar{x} \in \mathcal{R}(\bar{p})$ be a local minimizer of the problem $(\mathcal{P}_{\bar{p}})$. Problem $(\mathcal{P}_{\bar{p}})$ is called *calm* at $\bar{x}$ if there exist positive real constants r and ζ such that

$$\varphi(x) \geq \varphi(\bar{x}) - \zeta d(p, \bar{p}), \quad \forall x \in \mathrm{B}(\bar{x}, r) \cap \mathcal{R}(p), \ \forall p \in \mathrm{B}(\bar{p}, r).$$

Appeared firstly in [8], since then the above property become a fundamental regularity condition pervading the study of the sensitivity behaviour of variational problems, in the presence of perturbations (see [32]). In the present context, the introduction of problem calmness allows one to avoid the assumption of Lipschitz continuity on the objective functional. The price to be paid for enlarging the class of problems, to which the penalization technique can be applied, consists in a regularity requirement on the feasible region of the problem class, formalized as follows.

Definition 3.10. A set-valued mapping $\Xi : P \rightrightarrows X$ between metric spaces is said to be *semiregular* at $\bar{p} \in P$, *uniformly over* $\Xi(\bar{p})$, if if there exist positive real constants r and κ such that

$$(3.5) \qquad \mathrm{dist}\left(\bar{p}, \Xi^{-1}(x)\right) \leq \kappa d(x, \bar{x}), \quad \forall x \in \mathrm{B}(\bar{x}, r), \ \forall \bar{x} \in \Xi(\bar{p}).$$

The property formulated in Definition 3.10 is an enhanced version of a regularity notion that, to the best of the author's knowledge, was introduced in [21]. The latter is known to correspond to the well-known Lipschitz lower semicontinuity property for the inverse mapping of Ξ.

Remark 3.11. It is readily seen that, whenever $\Xi : P \rightrightarrows X$ is semiregular at $\bar{p}$, uniformly over $\Xi(\bar{p})$, then it holds

$$(3.6) \qquad \text{dist}\left(\bar{p}, \Xi^{-1}(x)\right) \leq \kappa \text{dist}\left(x, \Xi(\bar{p})\right), \quad \forall x \in \text{B}(\Xi(\bar{p}), r/2).$$

Indeed, taken $x \in \text{B}(\Xi(\bar{p}), r/2) \backslash \Xi(\bar{p})$ and any $\epsilon \in (0, 1)$, there exists $\bar{x}_\epsilon \in \Xi(\bar{p})$ such that

$$d(x, \bar{x}_\epsilon) < (1 + \epsilon)\text{dist}\left(x, \Xi(\bar{p})\right) < r.$$

Then, inequality (3.5) applies, so it results in

$$\text{dist}\left(\bar{p}, \Xi^{-1}(x)\right) \leq \kappa d(x, \bar{x}_\epsilon) < \kappa(1 + \epsilon)\text{dist}\left(x, \Xi(\bar{p})\right),$$

whence inequality (3.6) follows by arbitrariness of $\epsilon \in (0, 1)$. Actually, the validity of (3.6) is an equivalent reformulation of the uniform semiregularity of Ξ at $\bar{p}$, as one checks immediately.

One is now in a position to establish the next result about exact penalization.

Theorem 3.12. *With reference to a parameterized family of problems* $(\mathcal{P}_p)$, *let* $\bar{x} \in \mathcal{R}(\bar{p})$ *be a local minimizer of problem* $(\mathcal{P}_{\bar{p}})$. *Suppose that:*

    (i) (X, d) *is metrically complete;*
    (ii) φ *is l.s.c. at* $\bar{x}$;
    (iii) $\mathcal{R} : P \rightrightarrows X$ *is semiregular at* $\bar{p}$, *uniformly over* $\mathcal{R}(\bar{p})$;
    (iv) *problem* $(\mathcal{P}_{\bar{p}})$ *is calm at* $\bar{x}$;

and that there exists $\delta > 0$ *such that:*

    (v) $\Psi(p, \cdot) : X \rightrightarrows Y$ *is Hausdorff u.s.c. and set-covering on* X *with constant* $\alpha_p > 0$, *for each* $p \in \text{B}(\bar{p}, \delta)$;
    (vi) $\Phi(p, \cdot) : X \longrightarrow \mathcal{B}(Y)$ *is Lipschitz on* X *with constant* $\beta_p \in (0, \alpha_p)$, *for each* $p \in \text{B}(\bar{p}, \delta)$.

Then, there exists $l > 0$ *such that the penalty functional* $\varphi_l(\bar{p}, \cdot)$ *is exact at* $\bar{x}$.

Proof. Let us start with noting that, under the current assumptions, by virtue of Theorem 3.3 it is $\text{dom}\,\mathcal{R} \supseteq \text{B}(\bar{p}, \delta)$ and the following estimate holds

$$(3.7) \qquad \text{dist}\left(x, \mathcal{R}(p)\right) \leq \frac{\text{exc}\left(\Phi(p, x), \Psi(p, x)\right)}{\alpha_p - \beta_p}, \quad \forall x \in X, \ \forall p \in \text{B}(\bar{p}, \delta).$$

Recall that, according to what has been noticed in Remark 3.2, the mapping $\mathcal{R}$ is closed valued. By hypothesis (iii), in the light of Remark 3.11, there exist $r > 0$ and $\kappa > 0$ such that

$$\text{dist}\left(\bar{p}, \mathcal{R}^{-1}(x)\right) \leq \kappa \text{dist}\left(x, \mathcal{R}(\bar{p})\right), \quad \forall x \in \text{B}(\mathcal{R}(\bar{p}), r).$$

From the last inequality, on the account of the estimate (3.7), it follows

$$(3.8) \qquad \text{dist}\left(\bar{p}, \mathcal{R}^{-1}(x)\right) \leq \frac{\kappa}{\alpha_{\bar{p}} - \beta_{\bar{p}}}\text{exc}\left(\Phi(\bar{p}, x), \Psi(\bar{p}, x)\right),$$

$$\forall x \in \text{B}(\mathcal{R}(\bar{p}), r).$$

Assume now, ab absurdo, that for each $l > 0$ the penalty functional $\varphi_l(\bar{p}, \cdot)$ fails to be exact at $\bar{x}$, namely for each $l > 0$ there exists $n \in \mathbb{N}$, with $n > l$, and $x_n \in \text{B}(\bar{x}, 1/n)$ such that

$$(3.9) \qquad \varphi(x_n) + n \cdot \text{exc}\left(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)\right) < \varphi(\bar{x}).$$

Since $\bar{x}$ is a local solution to problem $(\mathcal{P}_{\bar{p}})$, for each $n \in \mathbb{N}$ larger than a proper natural number it must be $x_n \notin \mathcal{R}(\bar{p})$, that is, as multifunctions take closed values, $\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)) > 0$. Moreover, by virtue of the lower semicontinuity of φ at $\bar{x}$ and of the fact that the sequence $(x_n)_{n \in \mathbb{N}}$ converges to $\bar{x}$ as $n \to \infty$, from inequality (3.9) one obtains

$$
\begin{aligned}
\limsup_{n \to \infty} n \cdot \mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)) &\leq \limsup_{n \to \infty}\,[\varphi(\bar{x}) - \varphi(x_n)] \\
&= \varphi(\bar{x}) - \liminf_{n \to \infty} \varphi(x_n) \leq 0.
\end{aligned}
$$

Then one deduces that

$$
(3.10) \qquad \lim_{n \to \infty}\,\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)) = 0.
$$

Since, as already observed, $x_n \longrightarrow \bar{x}$ as $n \to \infty$, it is possible to assume without loss of generality that $x_n \in \mathrm{B}(\bar{x}, r)$, and hence $x_n \in \mathrm{B}(\mathcal{R}(\bar{p}), r)$. This fact enables one to apply inequality (3.8), from which one obtains

$$
\mathrm{dist}\,\left(\bar{p}, \mathcal{R}^{-1}(x_n)\right) \leq \frac{\kappa}{\alpha_{\bar{p}} - \beta_{\bar{p}}}\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)).
$$

This means that, corresponding to a costant $\tilde{\kappa} > \kappa$, a suitable $p_n \in \mathcal{R}^{-1}(x_n)$ can be found, such that the inequality

$$
(3.11) \qquad d(p_n, \bar{p}) < \frac{\tilde{\kappa}}{\alpha_{\bar{p}} - \beta_{\bar{p}}}\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n))
$$

holds for every $n \in \mathbb{N}$ large enough. Observe that, as $x_n \in \mathcal{R}(p_n)$ and $x_n \notin \mathcal{R}(\bar{p})$, then it must be $p_n \neq \bar{p}$. By combining inequalities (3.11) and (3.9), one finds

$$
\frac{\tilde{\kappa}}{\alpha_{\bar{p}} - \beta_{\bar{p}}} \cdot \frac{\varphi(x_n) - \varphi(\bar{x})}{d(p_n, \bar{p})} \leq \frac{\varphi(x_n) - \varphi(\bar{x})}{\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n))} < -n,
$$

wherefrom, for $x_n \in \mathrm{B}(\bar{x}, r) \cap \mathcal{R}(p_n)$ and every $n \in \mathbb{N}$ large enough, one gets

$$
(3.12) \qquad \varphi(x_n) < \varphi(\bar{x}) - \frac{n(\alpha_{\bar{p}} - \beta_{\bar{p}})}{\tilde{\kappa}}d(p_n, \bar{p}).
$$

Notice that, owing to inequality (3.10), the estimate in (3.11) entails that $p_n \longrightarrow \bar{p}$ as $n \to \infty$. Consequently, inequality (3.12) contradicts the hypothesis (iv) about the calmness at $\bar{x}$ of problem $(\mathcal{P}_{\bar{p}})$. Thus, the proof is complete. $\qquad \square$

Remark 3.13. Among the hypotheses of Theorem 3.12, (iii) and (iv) are not directly formulated in terms of problem data. Conditions for mapping $\mathcal{R}$ to be uniformly regular at $\bar{p}$ can be derived by working, under proper assumptions on Φ and Ψ, the general characterization for the semiregularity of a mapping $\Xi : P \rightrightarrows X$, which is expressed by the positivity of the constant

$$
\vartheta[\Xi](\bar{p}, \bar{x}) = \liminf_{\substack{x \notin \Xi(\bar{p}) \\ x \to \bar{x}}} \frac{\mathrm{dist}\,(x, \Xi(\bar{p}))}{\mathrm{dist}\,(\bar{p}, \Xi^{-1}(x))}
$$

(see [21]). A sufficient condition for a parameterized problem $(\mathcal{P}_{\bar{p}})$ to be calm can be expressed in terms of calmness from below of the related value function $\nu : P \longrightarrow \mathbb{R} \cup \{\pm\infty\}$, defined as

$$
\nu(p) = \inf_{x \in \mathcal{R}(p)} \varphi(x).
$$

More precisely, if $\bar{x} \in \mathcal{R}(\bar{p})$ is a global solution to problem $(\mathcal{P}_{\bar{p}})$, then $(\mathcal{P}_{\bar{p}})$ is calm $\bar{x}$ povided that ν is calm from below at $\bar{p}$, i.e.

$$\liminf_{p \to \bar{p}} \frac{\nu(p) - \nu(\bar{p})}{d(p, \bar{p})} > -\infty$$

(see, for more details, [34]).

Theorem 3.12 provides a sufficient condition for the exactness of the penalty functional, where problem calmness plays a crucial role. In order to enlighten the intriguing connection between these two properties, this subsection is concluded by a proposition that, in the setting under examination, singles out certain conditions upon which from exact penalization it is possible to derive problem calmness.

Proposition 3.14. *With reference to a parameterized family of problems $(\mathcal{P}_p)$, let $\bar{x} \in \mathcal{R}(\bar{p})$ be a local minimizer of $(\mathcal{P}_{\bar{p}})$. Suppose that:*

   (i) *mapping $\Phi : P \times X \longrightarrow \mathcal{B}(Y)$ is locally Lipschitz around $(\bar{p}, \bar{x})$;*
   (ii) *mapping Ψ is partially Lipschitz u.s.c. at $\bar{p}$, uniformly in x, i.e. there exist $r > 0$ and $\zeta > 0$ such that*

$$\Psi(p, x) \subseteq \mathrm{B}(\Psi(\bar{p}, x), \zeta d(p, \bar{p})), \quad \forall p \in \mathrm{B}(\bar{p}, \zeta), \ \forall x \in \mathrm{B}(\bar{x}, \zeta);$$

   (iii) *there exists $l > 0$ such that $\varphi_l(\bar{p}, \cdot)$ is an exact penalty functional.*

Then, problem $(\mathcal{P}_{\bar{p}})$ is calm at $\bar{x}$.

Proof. The thesis can be proved again by a reductio ad absurdum. So assume $(\mathcal{P}_{\bar{p}})$ to be not calm at $\bar{x}$. This amounts to say that for each $n \in \mathbb{N}$ it is possible to find $p_n \in \mathrm{B}(\bar{p}, 1/n) \backslash \{\bar{p}\}$ and $x_n \in \mathcal{R}(p_n) \cap \mathrm{B}(\bar{x}, 1/n)$ such that

$$(3.13) \qquad \varphi(x_n) < \varphi(\bar{x}) - nd(p_n, \bar{p}).$$

By hypothesis (ii), since $p_n \to \bar{p}$ and $x_n \to \bar{x}$ as $n \to \infty$, one has

$$(3.14) \qquad \mathrm{exc}\,(\Psi(p_n, x_n), \Psi(\bar{p}, x_n)) \leq \zeta d(p_n, \bar{p}),$$

for each $n \in \mathbb{N}$ large enough. On the other hand, since Φ is locally Lipschitz around $(\bar{p}, \bar{x})$, for some $\tau > 0$ it is true that

$$(3.15) \qquad \mathrm{exc}\,(\Phi(\bar{p}, x_n), \Phi(p_n, x_n)) \leq \tau d(p_n, \bar{p}),$$

for every $n \in \mathbb{N}$ large enough. By recalling that $x_n \in \mathcal{R}(p_n)$, so that $\Phi(p_n, x_n) \subseteq \Psi(p_n, x_n)$, from inequalities (3.15) and (3.14) one obtains

$$
\begin{aligned}
\mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)) \ &\leq \ \mathrm{exc}\,(\Phi(\bar{p}, x_n), \Phi(p_n, x_n)) \\
&+ \ \mathrm{exc}\,(\Phi(p_n, x_n), \Psi(p_n, x_n)) \\
&+ \ \mathrm{exc}\,(\Psi(p_n, x_n), \Psi(\bar{p}, x_n)) \\
&\leq \ (\tau + \zeta) d(p_n, \bar{p}).
\end{aligned}
$$

By combining the above estimate with inequality (3.13), one finds

$$\varphi(x_n) < \varphi(\bar{x}) - \frac{n}{\tau + \zeta} \mathrm{exc}\,(\Phi(\bar{p}, x_n), \Psi(\bar{p}, x_n)).$$

As the last inequality is true for each $n \in \mathbb{N}$ larger than a proper natural and for the corresponding $x_n \in \mathrm{B}(\bar{x}, 1/n)$, the hypothesis (iii) about the existence of

an exact penalty functional turns out to be contradicted. Thus, the argument by contradiction is complete.   □

Acknowledgements

Several valuable remarks and suggestions by two anonymous referees contributed to improve substantially the quality of this paper. The author would like to express his gratitude to both of them.

References

[1] A. V. Arutyunov, *Covering mappings in metric spaces, and fixed points*, Dokl. Akad. Nauk **416** (2007), 151–155 [in Russian].
[2] A. V. Arutyunov, E. R. Avakov, B. D. Gel'man, A. V. Dmitruk and V. V. Obukhovskii, *Locally covering maps in metric spaces and coincidence points*, J. Fixed Point Theory Appl. **5** (2009), 105–127.
[3] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser Boston, Inc., Boston, MA, 1990.
[4] E. R. Avakov, A. V. Arutyunov and A. F. Izmailov, On estimates for the rate of convergence of power penalty methods, Comput. Math. Math. Phys. **44** (2004), 1684–1695.
[5] E. R. Avakov, A. V. Arutyunov and A. F. Izmailov, Exact penalties for optimization problems with 2-regular equality constraints, Comput. Math. Math. Phys. **48** (2008), 346–353.
[6] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932 [in French].
[7] J. M. Borwein and D. M. Zhuang, *Verifiable necessary and sufficient conditions for openness and regularity for set-valued and single-valued maps* J. Math. Anal. Appl. **134** (1988), 441–459.
[8] F. H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res. **1** (1976), 165–174.
[9] A. V. Dmitruk, *On nonlocal metric regularity of nonlinear operators*, Control and Cybernetics **34** (2005), 723–746.
[10] A. V. Dmitruk, A. A. Milyutin and N. P. Osmolovskiĭ, *Lyusternik's theorem and the theory of extrema*, Russian Math. Surveys **35** (1980), 11–51.
[11] M. V. Dolgopolik, *A unifying theory of exactness of linear penalty functions*, forthcoming on Optimization, http://dx.doi.org/10.1080/02331934.2015.1122005.
[12] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings. A View from Variational Analysis*, Springer, New York, 2014.
[13] I. I. Eremin, *The method of penalties in convex programming*, Dokl. Akad. Nauk SSSR **173** (1967), 748–751.
[14] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vol. I, Springer-Verlag, New York, 2003.
[15] H. Frankowska, *An open mapping principle for set-valued maps*, J. Math. Anal. Appl. **127** (1987), 172–180.
[16] B. D. Gel'man and S. E. Zhukovskiy, *Covering maps of spaces of compact subsets*, Math. Notes **93** (2013), 539–544.
[17] L. M. Graves, *Some mapping theorems*, Duke Math. J. **17** (1950), 111–114.
[18] A. D. Ioffe, *Metric regularity and subdifferential calculus*, Russian Math. Surveys **55** (2000), 501–558.
[19] A. D. Ioffe, *Metric regularity, fixed points and some associated problems of variational analysis*, J. Fixed Point Theory Appl. **15** (2014), 67–99.
[20] J. Jahn, *Mathematical Vector Optimization in Partially Ordered Linear Spaces*, Verlag Peter D. Lang, Frankfurt am Main, 1986.
[21] A. Y. Kruger, *About stationarity and regularity in variational analysis*, Taiwanese J. Math. **13** (2009), 1737–1785.
[22] L. A. Lyusternik, *On the conditional extrema of functionals*, Mat. Sbornik **41** (1934), 390–401. [in Russian]

[23] B. S. Mordukhovich, *Nonsmooth analysis with nonconvex generalized differentials and conjugate mappings* Dokl. Akad. Nauk BSSR **28** (1984), 976–979. [in Russian]

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

[25] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I: Basic Theory*, Springer, Berlin, 2006.

[26] B. S. Mordukhovich and Y.H. Shao, *Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces*, Nonlinear Anal. **25** (1995), 1401–1424.

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

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

[29] S. M. Robinson, *Normed convex processes*, Trans. Amer. Math. Soc. **174** (1972), 127–140.

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

[31] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, N. J., 1970.

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

[33] S. Rolewicz, *Functional Analysis and Control Theory*, D. Reidel Publishing Co., Dordrecht; PWN—Polish Scientific Publishers, Warsaw, 1987.

[34] A. Uderzo, *Exact penalty functions and calmness for mathematical programming under nonlinear perturbations*, Nonlinear Anal. **73** (2010), 1596–1609.

[35] N. D. Yen, J.-C. Yao and B.T. Kien, *Covering properties at positive-order rates of multifunctions and some related topics*, J. Math. Anal. Appl. **338** (2008), 467–478.

[36] W. I. Zangwill, *Nonlinear programming via penalty functions*, Management Science **13** (1967), 344–358.

[37] A. J. Zaslavski, *Optimization on Metric and Normed Spaces*, Springer, New York, 2010.

A. Uderzo
University of Milano-Bicocca, Dept. of Mathematics and Applications, Milan, Italy
E-mail address: `amos.uderzo@unimib.it`