Pure and Applied Functional Analysis
Volume 4, Number 2, 2019, 427–445

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

AN EXTENSION OF THE POLYAK CONVEXITY PRINCIPLE WITH APPLICATION TO NONCONVEX OPTIMIZATION

AMOS UDERZO

Abstract. The main problem considered in the present paper is to single out classes of convex sets, whose convexity property is preserved under nonlinear smooth transformations. Extending an approach due to B.T. Polyak, the present study focusses on the class of uniformly convex subsets of Banach spaces. As a main result, a quantitative condition linking the modulus of convexity of such kind of set, the regularity behaviour around a point of a nonlinear mapping and the Lipschitz continuity of its derivative is established, which ensures the images of uniformly convex sets to remain uniformly convex. Applications of the resulting convexity principle to the existence of solutions, their characterization and to the Lagrangian duality theory in constrained nonconvex optimization are then discussed.

1. Introduction

In many fields of mathematics, persistence phenomena of specific geometrical properties under various kind of transformations have been often a subject of interest and study. Transformations, when possible formalized by mappings acting among spaces, sometimes have been classified on the basis of features in a structure that they can preserve (whence the very term “morphism”). Convexity is a geometrical property which emerged in ancient times, at the very beginning of geometry, and since then remained essentially unchanged for almost two millennia and half. This happened by virtue of the great variety of successful applications that it found in many different areas. In particular, the relevant role played by convexity in optimization and control theory is widely recognized. This led to develop a branch of mathematics, called convex analysis, that elected convexity as its main topic of study. In spite of such an interest and motivations, not much seems to be known up to now about phenomena of persistence of convexity under nonlinear transformations. Yet, advances in this direction would have a certain impact on the analysis of optimization problems. Historically, the first results somehow connected with the issue at the study relate to the numerical range of quadratic mappings (namely, mappings whose components are quadratic forms) and can be found in [7] (see also [21]). A notable step ahead was made when the preservation of convexity of small balls under smooth regular transformations between Hilbert spaces was established by B.T. Polyak (see [22]). After that, some other contributions to understanding

2010 *Mathematics Subject Classification.* 49J27, 49K27, 52A05, 90C26, 90C46, 90C48.
Key words and phrases. Modulus of convexity, uniformly convex sets, openness at a linear rate/metric regularity, constrained optimization, strong duality.

the phenomenon in a similar context were given by [1, 4, 10, 25]. Various applications of it to topics in linear algebra, optimization and control theory are presented in [21, 22, 23, 24, 25]

In the present paper, by following the approach introduced by B.T. Polyak, the study of classes of sets with persistent convexity properties is carried on. More precisely, the analysis here proposed focusses on the class of uniformly convex subsets of certain Banach spaces. An interest in similar classes of sets, in connection with the problem under study, appears already in [22], where strongly convex sets are actually mentioned. This seems to be rather natural, inasmuch as elements of such classes share the essential geometrical features of balls in a Hilbert space: nonempty interior, boundedness and, what plays a crucial role, a uniform rotundity, which implies a boundary consisting of extreme points only. The feature last mentioned is captured and quantitatively expressed by the notion of modulus of convexity of a set. In developing the Polyak's approach, the main idea behind the investigations exposed in the paper is that, if the modulus of convexity of a given set matches the smoothness and the regularity property of a given nonlinear mapping, then the persistence of convexity under that mapping can be guaranteed. The understanding of such a fundamental relation between quantitative aspects of the convexity property for a set and the quantitative regularity behaviour of a mapping acting on it should shed light on the general phenomenon under study. Concretely, this leads to enrich the class of sets interested by the phenomenon. In turn, since the persistence of convexity under nonlinear transformations is at the origin of a certain qualification (in terms of solution existence and characterization) observed in optimization problems with possibly nonconvex data, the result here established allows one to enlarge the class of problems for which the consequent benefits can be expected.

The contents of the paper are arranged in the next sections as follows. In Section 2, the notion of modulus of convexity of a set and of uniformly convexity are recalled, along with several examples and related facts, useful for the subsequent analysis. Besides, the regularity behaviour of a nonlinear smooth mapping, namely its openness at a linear rate, is entered as a crucial tool, along with the related exact bound. In Section 3, the main result of the paper, which is an extension of the aforementioned convexity principle due to B.T. Polyak, is established and some of its features are discussed. In Section 4, some applications of the main result to nonconvex constrained optimization problems are provided.

2. Notations and preliminaries

The basic notations in use throughout the paper are as follows. $\mathbb{R}$ denotes the real number set. Given a metric space (X,d), an element $x_0 \in X$ and $r \geq 0$, $\mathrm{B}(x_0,r) = \{x \in X :\ d(x,x_0) \leq r\}$ denotes the (closed) ball with center x_0 and radius r. In particular, in a Banach space, the unit ball centered at the null vector will be indicated by $\mathbb{B}$, whereas the unit sphere by $\mathbb{S}$. The distance of $x_0 \in X$ from a set $S \subseteq X$ is denoted by $\mathrm{dist}(x_0,S)$. If $S \subseteq X$, $\mathrm{B}(S,r) = \{x \in X :\ \mathrm{dist}(x,S) \leq r\}$ denotes the (closed) r-enlargement of S. The diameter of a set $S \subseteq X$ is defined as $\mathrm{diam}\, S = \sup\{d(x_1,x_2) :\ x_1,\, x_2 \in S\}$. By $\mathrm{int}\, S$, $\mathrm{cl}\, S$ and $\mathrm{bd}\, S$ the topological interior, the closure and the boundary of a set S are marked, respectively. If S is a subset of a Banach space $(\mathbb{X}, \|\cdot\|)$, $\mathrm{ext}\, S$ denotes the set of all extreme points of S, in the sense of convex analysis, $\mathbf{0}$ stands for the null element of $\mathbb{X}$ and $[x_1, x_2]$ denotes the closed line segment with endpoints $x_1,\, x_2 \in \mathbb{X}$. Given a function $h : X \longrightarrow Y$ between metric spaces and a set $U \subseteq X$, h is said to be Lipschitz continuous on U if there exists a constant $\ell > 0$ such that

$$d(h(x_1), h(x_2)) \leq \ell d(x_1, x_2), \quad \forall x_1,\, x_2 \in U. \tag{2.1}$$

The infimum over all values ℓ making the last inequality satisfied on U is called exact bound of Lipschitz continuity of h on U and is denoted by $\mathrm{lip}(h, U)$, i.e.

$$\mathrm{lip}(h, U) = \inf\{\ell \geq 0 : \quad \text{inequality (2.1) holds}\}.$$

The Banach space of all bounded linear operators between the Banach spaces $\mathbb{X}$ and $\mathbb{Y}$, equipped with the operator norm, is denoted by $(\mathcal{L}(\mathbb{X}, \mathbb{Y}), \|\cdot\|_{\mathcal{L}})$. If, in particular, it is $\mathbb{Y} = \mathbb{R}$, the simpler notation $(\mathbb{X}^*, \|\cdot\|_*)$ is used. The null vector in a dual space is marked by $\mathbf{0}^*$, whereas the unit sphere by $\mathbb{S}^*$, with $\langle\cdot,\cdot\rangle$ marking the duality pairing a space and its dual. Given a mapping $f : \Omega \longrightarrow \mathbb{Y}$, with Ω open subset of $\mathbb{X}$, and $x_0 \in \Omega$, the Gatêaux derivative of f at x_0 is denoted by $\mathrm{D}f(x_0)$. If f is Gatêaux differentiable at each point of Ω and the mapping $\mathrm{D}f : \Omega \longrightarrow \mathcal{L}(\mathbb{X}, \mathbb{Y})$ is Lipschitz continuous on Ω, f is said to be of class $\mathrm{C}^{1,1}(\Omega)$.

Remark 2.1. (i) In view of a subsequent employment, let us recall that, whenever $f : \Omega \longrightarrow \mathbb{Y}$ is a mapping of class $\mathrm{C}^{1,1}(\Omega)$ between Banach spaces, with Ω open subset of $\mathbb{X}$ and $x_1,\, x_2 \in \Omega$ are such that $[x_1, x_2] \subseteq \Omega$, the following estimate holds true (see, for instance, [26, Lemma 2.7])

$$\left\| \frac{f(x_1) + f(x_2)}{2} - f\left(\frac{x_1 + x_2}{2}\right) \right\| \leq \frac{\mathrm{lip}(\mathrm{D}f, \Omega)}{8} \|x_1 - x_2\|^2, \tag{2.2}$$

where $\mathrm{lip}(\mathrm{D}f, \Omega)$ denotes the exact bound of Lipschitz continuity of $\mathrm{D}f$ on Ω.

(ii) It is not difficult to see that, if $S \subseteq \Omega$ is a bounded set, i.e. $\operatorname{diam} S < +\infty$, and $f \in \mathrm{C}^{1,1}(\Omega)$, then it must be

$$\sup_{x \in S} \|\mathrm{D}f(x)\|_{\mathcal{L}} < +\infty.$$

Furthermore, if in addition S is convex, then letting $\beta_S = \sup_{x \in S} \|\mathrm{D}f(x)\|_{\mathcal{L}}$, as an immediate consequence of the mean-value theorem, one obtains

$$\operatorname{diam} f(S) \leq \beta_S \operatorname{diam} S,$$

that is $f(S)$ is bounded too.

2.1. **Uniformly convex sets.**

Definition 2.2. (i) Let $S \subseteq \mathbb{X}$ be a nonempty, closed and convex subset of a real Banach space. The function $\delta_S : [0, \operatorname{diam} S) \longrightarrow [0, +\infty)$ defined by

$$\delta_S(\epsilon) = \sup\left\{ \delta \geq 0 : \; \mathrm{B}\left(\frac{x_1 + x_2}{2}, \delta\right) \subseteq S, \; \forall x_1,\, x_2 \in S : \; \|x_1 - x_2\| = \epsilon \right\}$$

is called *modulus of convexity* of the set S. Whenever the value of $\operatorname{diam} S$ is attained at some pair $x_1,\, x_2 \in S$, the function δ_S will be meant to be naturally extended to $[0, \operatorname{diam} S]$.

(ii) After [20], a nonempty, closed and convex set $S \subseteq \mathbb{X}$, with $S \neq \mathbb{X}$, is said to be *uniformly convex* provided that

$$\delta_S(\epsilon) > 0, \quad \forall \epsilon \in \begin{cases} (0, \operatorname{diam} S], & \text{if } \operatorname{diam} S \text{ is attained on } S, \\ (0, \operatorname{diam} S), & \text{otherwise.} \end{cases}$$

Since $\operatorname{diam} S$ vanishes if S is a singleton, Definition 2.2 (ii) does not exclude such kind of convex sets. Nevertheless, as singletons are of minor interest in connection with the problem at the issue, henceforth a uniformly convex set will be always assumed to contain at least two distinct points.

Example 2.3. (i) Balls in a uniformly convex Banach space may be viewed as a paradigma for the notion of uniform convexity for sets. Recall that, after [5], a Banach space $(\mathbb{X}, \|\cdot\|)$ is said to be *uniformly convex* (or to have a uniformly convex norm) if

$$\delta_{\mathbb{X}}(\epsilon) = \inf \left\{ 1 - \left\| \frac{x_1 + x_2}{2} \right\| : \ x_1, \, x_2 \in \mathbb{B}, \ \|x_1 - x_2\| = \epsilon \right\} > 0, \forall \epsilon \in (0, 2].$$

The function $\delta_{\mathbb{X}}$ is called modulus of convexity of the space $(\mathbb{X}, \|\cdot\|)$. In fact, it is possible to prove that

$$\delta_{\mathbb{B}}(\epsilon) = \delta_{\mathbb{X}}(\epsilon), \quad \forall \epsilon \in (0, 2].$$

Such classes of Banach spaces as l^p and L^p, with $1 < p < \infty$, are known to consist of uniformly convex spaces. In particular, every Hilbert space is uniformly convex. Since every uniformly convex Banach space must be reflexive (according to the Milman-Pettis Theorem), the spaces l^1, L^1, L^∞, $C([0, 1])$ and c_0 fail to be. For $p \geq 2$, the exact expression of the modulus of convexity of the spaces l^p and L^p is given by

$$\delta_{l^p}(\epsilon) = \delta_{L^p}(\epsilon) = 1 - \left[1 - \left(\frac{\epsilon}{2} \right)^p \right]^{1/p}, \quad \forall \epsilon \in (0, 2].$$

For more details on uniformly convex Banach spaces and properties of their moduli the reader may refer to [6, 12, 16]. A useful remark enlightening the connection between the notions of uniform convexity for sets and uniform convexity of Banach spaces can be found in [2, Theorem 2.3]: a Banach space can contain a closed uniformly convex set iff it admits an equivalent uniformly convex norm. Such class of Banach spaces have been characterized in terms of superreflexivity in [11]. Throughout the present paper, the Banach space $(\mathbb{X}, \|\cdot\|)$ will be supposed to be equipped with a uniformly convex norm.

(ii) After [18, 19], given a positive real r, a subset $S \subseteq \mathbb{X}$ of a Banach space is said to be *r-convex* (or *strongly convex* of radius r) if there exists $M \subseteq \mathbb{X}$, with $M \neq \mathbb{X}$, such that

$$S = \bigcap_{x \in M} \mathrm{B}\,(x, r) \neq \varnothing.$$

It is readily seen that, if a Banach space $(\mathbb{X}, \|\cdot\|)$ is uniformly convex with modulus $\delta_{\mathbb{X}}$, then any strongly convex set $S \subseteq \mathbb{X}$ with radius r is uniformly convex and its modulus of convexity satisfies the relation

$$\delta_S(\epsilon) \geq r \delta_{\mathbb{X}} \left(\frac{\epsilon}{r} \right), \quad \forall \epsilon \in (0, \operatorname{diam} S). \tag{2.3}$$

(iii) Let $\theta : [0, +\infty) \longrightarrow [0, +\infty)$ be an increasing function vanishing only at 0. Recall that, according to [27], a function $\varphi : \mathbb{X} \longrightarrow \mathbb{R}$ is said to be *uniformly convex with modulus* θ if it holds

$$\varphi(tx_1 + (1-t)x_2) \leq t\varphi(x_1) + (1-t)\varphi(x_2) - t(1-t)\theta(\|x_1 - x_2\|), \quad \forall x_1,\, x_2 \in \mathbb{X},\ \forall t \in [0,1].$$

If, in particular, it is $\theta(s) = \kappa s^2$, a uniformly convex function with such a modulus is called *strongly convex*. Sublevel sets of Lipschitz continuous uniformly convex functions are uniformly convex sets. More precisely, given $\alpha > 0$, if φ is Lipschitz continuous on $\mathbb{X}$, with exact bound $\mathrm{lip}(\varphi, \mathbb{X}) > 0$, then the set $[\varphi \leq \alpha] = \{x \in \mathbb{X} : \varphi(x) \leq \alpha\}$ turns out to be uniformly convex with modulus

$$\delta_{[\varphi\leq\alpha]}(\epsilon) \geq \frac{\theta(\epsilon)}{4\mathrm{lip}(\varphi, \mathbb{X})}, \quad \forall \epsilon \in (0, \mathrm{diam}\,[\varphi \leq \alpha]). \tag{2.4}$$

Indeed, fixed $\epsilon \in (0, \mathrm{diam}\,[\varphi \leq \alpha])$, take $x_1,\, x_2 \in [\varphi \leq \alpha]$, with $x_1 \neq x_2$ and $\|x_1 - x_2\| = \epsilon$, and set $\bar{x} = \frac{1}{2}(x_1 + x_2)$. By the uniform convexity of φ with modulus θ one has

$$\varphi(\bar{x}) \leq \frac{\varphi(x_1) + \varphi(x_2)}{2} - \frac{\theta(\|x_1 - x_2\|)}{4}.$$

Therefore, for an arbitrary $\eta > 0$, by the Lipschitz continuity of φ on $\mathbb{X}$, one finds

$$\begin{aligned} \varphi(x) &= \varphi(x) - \varphi(\bar{x}) + \varphi(\bar{x}) \\ &\leq (\mathrm{lip}(\varphi, \mathbb{X}) + \eta)\frac{\theta(\epsilon)}{4(\mathrm{lip}(\varphi, \mathbb{X}) + \eta)} + \alpha - \frac{\theta(\epsilon)}{4} \leq \alpha, \end{aligned}$$

for every $x \in \mathrm{B}\left(\bar{x}, \frac{\theta(\epsilon)}{4(\mathrm{lip}(\varphi,\mathbb{X})+\eta)}\right)$. Thus, it results in

$$\mathrm{B}\left(\bar{x}, \frac{\theta(\epsilon)}{4(\mathrm{lip}(\varphi, \mathbb{X}) + \eta)}\right) \subseteq [\varphi \leq \alpha],$$

so

$$\delta_{[\varphi\leq\alpha]}(\epsilon) \geq \frac{\theta(\epsilon)}{4(\mathrm{lip}(\varphi, \mathbb{X}) + \eta)}.$$

The estimate in (2.4) follows by arbitrariness of η.

It is not difficult to see that, given two subsets S_1 and S_2 of $\mathbb{X}$, it is $\delta_{S_1\cap S_2} \geq \min\{\delta_{S_1},\, \delta_{S_2}\}$. Therefore, the class of uniformly convex sets is closed under finite intersection. In contrast, unlike the class of convex sets, this class fails to be closed with respect to the Cartesian product. It is worth noting that, as the intersection of balls may yield a boundary with corners or a nonsmooth description, uniformly convex sets may exhibit such kind of pathology.

In the next remark, some known facts about uniformly convex sets are collected, which will be relevant to the subsequent analysis.

Remark 2.4. (i) Every uniformly convex set, which does not coincide with the entire space, is bounded (see [2]).

(ii) Directly from Definition 2.2, it follows that every uniformly convex set has nonempty interior. This fact entails that, while uniformly convex subsets are compact if living in finite-dimensional spaces, they can not be so in infinite-dimensional Banach spaces.

(iii) As a consequence of Definition 2.2, if any uniformly convex set S admits a modulus of convexity of power type 2, i.e. such that

$$\delta_S(\epsilon) \geq c\epsilon^2, \quad \forall \epsilon \in (0, \operatorname{diam} S), \tag{2.5}$$

for some $c > 0$, then it fulfils the following property: for every $\tilde{c} \in (0, c)$ it holds

$$\mathrm{B}\left(\frac{x_1 + x_2}{2}, \tilde{c}\|x_1 - x_2\|^2\right) \subseteq S, \quad \forall x_1,\, x_2 \in S.$$

It is worth noting that this happens for the balls in any Hilbert space or in the Banach spaces l^p and L^p, with $1 < p < 2$, where the following estimate is known to hold

$$\delta_{l^p}(\epsilon) = \delta_{L^p}(\epsilon) > \frac{p-1}{8}\epsilon^2, \quad \forall \epsilon \in (0, 2]$$

(see, for instance, [16]). Such a subclass of uniformly convex sets will play a prominent role in the main result of the paper.

(iv) For every uniformly convex set S, a constant $\beta > 0$ can be proved to exist such that

$$\delta_S(\epsilon) \leq \beta\epsilon^2, \quad \forall \epsilon \in (0, \operatorname{diam} S)$$

(see [2]). Thus, a modulus of convexity of the power 2 is a maximal one.

The next proposition provides a complete characterization of uniform convexity for subsets of a finite-dimensional Euclidean space in terms of extremality of their boundary points. Below, a variational proof of this fact is provided.

Proposition 2.5. *A convex compact subset $S \subseteq \mathbb{R}^n$, with nonempty interior, is uniformly convex iff* $\operatorname{ext} S = \operatorname{bd} S$.

Proof. Observe that by compactness of S, it is $\operatorname{bd} S \neq \varnothing$. Actually, the Krein-Milman theorem ensures that $\operatorname{ext} S \neq \varnothing$ also. Clearly, it is $\operatorname{ext} S \subseteq \operatorname{bd} S$. To begin with, assume that S is uniformly convex. Take any $\bar{x} \in \operatorname{bd} S$. If it were $\bar{x} \notin \operatorname{ext} S$, then there would exist $x_1,\, x_2 \in S\backslash\{\bar{x}\}$, with $x_1 \neq x_2$, such that $\bar{x} = \frac{x_1+x_2}{2}$. Observe that, as $\bar{x} \in \operatorname{bd} S$, the inclusion $\mathrm{B}(\bar{x}, \delta) \subseteq S$ can be true only for $\delta = 0$. Thus $\delta_S(\|x_1 - x_2\|) = 0$, contradicting the fact that S is uniformy convex.

Conversely, assume that the equality $\operatorname{ext} S = \operatorname{bd} S$ holds true. Fix an arbitrary $\epsilon \in (0, \operatorname{diam} S]$ (under the current hypotheses the value $\operatorname{diam} S$ is attained on S). Notice that, since S is compact, the set

$$S^2_\epsilon = \{(x_1, x_2) \in S \times S :\ \|x_1 - x_2\| = \epsilon\}$$

is still compact. Define the function $\vartheta : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow [0, +\infty)$ by setting

$$\vartheta(x_1, x_2) = \operatorname{dist}\left(\frac{x_1 + x_2}{2}, \mathbb{R}^n \backslash \operatorname{int} S\right).$$

Since such a function is continuous on $\mathbb{R}^n \times \mathbb{R}^n$, it attains its global minimum over S^2_ϵ at some point $(\hat{x}_1, \hat{x}_2) \in S^2_\epsilon$, with $\hat{x}_1 \neq \hat{x}_2$ as $\|\hat{x}_1 - \hat{x}_2\| = \epsilon$. If it were $\vartheta(\hat{x}_1, \hat{x}_2) = 0$, then it would happen that

$$\frac{\hat{x}_1 + \hat{x}_2}{2} \in \operatorname{bd} S.$$

The last inclusion contradicts the fact that $\frac{\hat{x}_1+\hat{x}_2}{2}$ is an extreme point of S. Therefore, one deduces that $\vartheta(\hat{x}_1, \hat{x}_2) > 0$. As it is true that

$$\delta_S(\epsilon) = \min_{(x_1,x_2)\in S^2_\epsilon} \vartheta(x_1, x_2) > 0,$$

the requirement in Definition 2.2 (ii) turns out to be satisfied. The arbitrariness of $\epsilon \in (0, \operatorname{diam} S]$ completes the proof. □

Proposition 2.5 can not be extended to infinite-dimensional spaces, where balls with $\operatorname{ext}\mathbb{B} = \operatorname{bd}\mathbb{B}$ can exist, yet failing to be uniformly convex (see [6]).

2.2. **Openness at a linear rate.** In the next definition, some notions and related results are recalled, which describe quantitatively a certain surjective behaviour of a mapping. Such a local property, in a synergical interplay with other features ($\mathrm{C}^{1,1}$-smoothness and uniform convexity) of the involved objects, allows one to achieve the main result in the paper.

Definition 2.6. Let $f : X \longrightarrow Y$ be a mapping between two metric spaces and $x_0 \in X$. The mapping f is said to be *open at a linear rate around* x_0 if there exist positive reals δ, ζ and σ such that

$$f(\mathrm{B}(x, r)) \supseteq \mathrm{B}(f(x), \sigma r) \cap \mathrm{B}(f(x_0), \zeta), \quad \forall x \in \mathrm{B}(x_0, \delta), \ \forall r \in [0, \delta]. \tag{2.6}$$

The role of a surjection property in preserving convexity of sets should not come as a surprise: the convexity of the image requires indeed line segments joining points in the image of a set to belong to the image, that is a certain openness/covering behaviour of the reference mapping.

It is well known (see, for instance, [9, 14, 17]) that the property of openness at a linear rate for a mapping f around x_0 can be equivalently reformulated as follows: there exist positive reals δ and κ such that

$$\operatorname{dist}\left(x, f^{-1}(y)\right) \le \kappa d(y, f(x)), \quad \forall x \in \mathrm{B}(x_0, \delta), \ \forall y \in \mathrm{B}(f(x_0), \delta). \tag{2.7}$$

Whenever the inequality (2.7) holds, f is said to be *metrically regular* around x_0. The infimum over all values κ for which there exists $\delta > 0$ such that (2.7) holds true is called *exact regularity bound* of f around x_0 and it will be denoted by $\operatorname{reg}(f, x_0)$, with the convention that $\operatorname{reg}(f, x_0) = +\infty$ means that f fails to be metrically regular around x_0.

Remark 2.7. (i) It is convenient to note that, whenever f is continuous at x_0, the inclusion defining the openness of f at a linear rate around x_0 takes the simpler form: there exists positive δ and σ such that

$$f(\mathrm{B}(x, r)) \supseteq \mathrm{B}(f(x), \sigma r), \quad \forall x \in \mathrm{B}(x_0, \delta), \ \forall r \in [0, \delta]. \tag{2.8}$$

(ii) From the inclusion (2.8) it is clear that, whenever a mapping f is open at a linear rate around x_0 and continuous at the same point, it holds

$$f(\operatorname{int} S) \subseteq \operatorname{int} f(S), \tag{2.9}$$

provided that $S \subseteq \mathrm{B}(x, \delta)$, where δ is as above. Indeed, if it is $x \in \operatorname{int} S$, then for some $r \in (0, \delta)$ it must be $\mathrm{B}(x, r) \subseteq S$. Therefore, one gets

$$\mathrm{B}(f(x), \sigma r) \subseteq f(\mathrm{B}(x, r)) \subseteq f(S).$$

In turn, from the inclusion (2.9), one deduces

$$f^{-1}(y) \cap S \subseteq \operatorname{bd} S, \quad \forall y \in \operatorname{bd} f(S).$$

As the behaviour formalized by openness at a linear rate/metric regularity plays a crucial role in a variety of topics in variational analysis, it has been widely investigated in the past decades and several criteria for detecting the occurrence of it are now at disposal. In the case of smooth mappings between Banach spaces, the main criterion for openness at a linear rate/metric regularity, known under the name of Lyusternik-Graves theorem, can be stated as follows (see [9, 14, 17]).

Theorem 2.8 (Lyusternik-Graves)**.** *Let $f : \mathbb{X} \longrightarrow \mathbb{Y}$ be a mapping between Banach spaces. Suppose that f is strictly differentiable at $x_0 \in \mathbb{X}$. Then, f is open at a linear rate around x_0 iff $\mathrm{D}f(x_0)$ is onto, i.e. $\mathrm{D}f(x_0)(\mathbb{X}) = \mathbb{Y}$.*

The above criterion is usually complemented with the following (primal and dual) estimates of the exact regularity bound, which are relevant for the present analysis:

$$\operatorname{reg}(f, x_0) = \sup_{\|y\| \leq 1} \inf\{\|x\| : \ x \in \mathrm{D}f(x_0)^{-1}(y)\}$$

and

$$\operatorname{reg}(f, x_0) = \left(\inf_{\|y^*\|_* = 1} \|\mathrm{D}f(x_0)^* y^*\|_* \right)^{-1} = \left(\operatorname{dist}\left(\mathbf{0}^*, \mathrm{D}f(x_0)^*(\mathbb{S}^*)\right)\right)^{-1},$$

where $\Lambda^* \in \mathcal{L}(\mathbb{Y}^*, \mathbb{X}^*)$ denotes the adjoint operator to $\Lambda \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ and the conventions

$$\inf \varnothing = +\infty \qquad \text{and} \qquad 1/0 = +\infty$$

are adopted. Remember that $\Lambda \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ is onto iff Λ^* has bounded inverse. It is worth noting that, when both $\mathbb{X}$ and $\mathbb{Y}$ are finite-dimensional Euclidean spaces, the condition on $\mathrm{D}f(x_0)$ to be onto reduces to the fact that Jacobian matrix of f at x_0 is full-rank. Furthermore, whenever $\mathrm{D}f(x_0)$ happens to be invertible, one has $\operatorname{reg}(f, x_0) = \|\mathrm{D}f(x_0)^{-1}\|_{\mathcal{L}}$.

3. An extension of the Polyak convexity principle

Given $c > 0$, let us introduce the following subclasses of uniformly convex subsets of $\mathbb{X}$, with modulus of convexity of power type 2:

$$\mathcal{UC}^2_c(\mathbb{X}) = \{S \subseteq \mathbb{X} : \ \delta_S(\epsilon) \geq c\epsilon^2, \ \forall \epsilon \in (0, \operatorname{diam} S)\}$$

and

$$\mathcal{UC}^2(\mathbb{X}) = \bigcup_{c>0} \mathcal{UC}^2_c(\mathbb{X}).$$

Remark 3.1. In the proof of the next theorem the following fact, which can be easily proved by an iterative bisection procedure, will be used: any closed subset V of a Banach space is convex iff $\frac{y_1+y_2}{2} \in V$, whenever y_1, $y_2 \in V$. It is easy to see that if V is not closed, this mid-point property does not imply the convexity of V. Consider, for instance, the set V defined by

$$V = \bigcup_{k=0}^{\infty} \left\{ \frac{i}{2^k} : \ i \in \{0, 1, 2, 3, \dots, 2^k\} \right\} \subseteq [0, 1].$$

Since V is countable, as a countable union of finite sets, it is strictly included in $[0, 1]$. Therefore V can not be convex, because it contains 0 and 1, even though it has the mid-point property, as one checks without difficulty.

Below, the main result of the paper is established.

Theorem 3.2. *Let $f : \Omega \longrightarrow \mathbb{Y}$ be a mapping between Banach spaces, with Ω open nonempty subset of $\mathbb{X}$. Let $x_0 \in \Omega$ and $c > 0$ such that:*

(i) $f \in \mathrm{C}^{1,1}(\mathrm{int}\,\mathrm{B}\,(x_0, r_0))$, *for some $r_0 > 0$;*
(ii) $\mathrm{D}f(x_0)$ *is onto;*
(iii) *it holds*

$$\frac{\mathrm{reg}(f, x_0) \cdot \mathrm{lip}(\mathrm{D}f, \mathrm{int}\,\mathrm{B}\,(x_0, r_0))}{8} < c.$$

Then, there exists $\rho \in (0, r_0)$ such that, for every $S \in \mathcal{UC}^2_c(\mathbb{X})$, with $S \subseteq \mathrm{B}\,(x_0, \rho)$ and $f(S)$ closed, it is $f(S) \in \mathcal{UC}^2(\mathbb{Y})$.

Proof. The proof is divided into two parts.
First part: Let us show that $f(S)$ is convex. According to the hypothesis (iii), it is possible to fix positive reals κ and ℓ in such a way that $\kappa > \mathrm{reg}(f, x_0)$, $\ell > \mathrm{lip}(\mathrm{D}f, \mathrm{int}\,\mathrm{B}\,(x_0, r_0))$, and the following inequality is fulfilled

$$\frac{\kappa\ell}{8} < c. \tag{3.1}$$

By virtue of hypotheses (i) and (ii), as f is in particular strictly differentiable at x_0, it is possible to invoke the Lyusternik-Graves theorem, ensuring that f is metrically regular around x_0. This means that there exist positive reals $\tilde{\kappa}$ and $\tilde{r}$ such that

$$\mathrm{reg}(f, x_0) < \tilde{\kappa} < \kappa, \qquad \tilde{r} \in (0, r_0),$$

and

$$\mathrm{dist}\left(x, f^{-1}(y)\right) \leq \tilde{\kappa}\|y - f(x)\|, \quad \forall x \in \mathrm{B}\,(x_0, \tilde{r})\,,\ \forall y \in \mathrm{B}\,(f(x_0), \tilde{r})\,. \tag{3.2}$$

Besides, by the continuity of f at x_0, corresponding to $\tilde{r}$ there exists $r_* \in (0, r_0)$ such that

$$f(x) \in \mathrm{B}\,(f(x_0), \tilde{r})\,, \quad \forall x \in \mathrm{B}\,(x_0, r_*)\,.$$

Then, take $\rho \in (0, \min\{\tilde{r}, r_*\})$. Notice that, in the light of Remark 2.7, up to a further reduction in the value of ρ, one can assume that for some $\sigma > 0$ it holds

$$f(\mathrm{B}\,(x, r)) \supseteq \mathrm{B}\,(f(x), \sigma r)\,, \quad \forall x \in \mathrm{B}\,(x_0, \rho)\,,\ \forall r \in [0, \rho]. \tag{3.3}$$

Now, take an arbitrary element $S \in \mathcal{UC}^2_c(\mathbb{X})$, with $S \subseteq \mathrm{B}\,(x_0, \rho)$ and such that $f(S)$ is closed. According to Remark 3.1, the convexity of $f(S)$ can be proved by showing that for every $y_1, y_2 \in f(S)$, with $y_1 \neq y_2$, it holds $\frac{y_1+y_2}{2} \in f(S)$. To this aim, let $x_1, x_2 \in S$ be such that $y_1 = f(x_1)$ and $y_2 = f(x_2)$. For convenience, set

$$\bar{x} = \frac{x_1 + x_2}{2} \qquad \text{and} \qquad \bar{y} = \frac{y_1 + y_2}{2}$$

Notice that, as $y_1 \neq y_2$, it must be also $x_1 \neq x_2$. Moreover, as $S \subseteq \mathrm{B}\,(x_0, \rho) \subseteq \mathrm{B}\,(x_0, r_*)$, one has $y_1, y_2 \in \mathrm{B}\,(f(x_0), \tilde{r})$ and therefore, by the convexity of a ball, one has also $\bar{y} \in \mathrm{B}(f(x_0), \tilde{r})$. Thus, since $\bar{x} \in \mathrm{B}(x_0, \tilde{r})$ and $y \in \mathrm{B}(f(x_0), \tilde{r})$, then inequality (3.2) implies

$$\text{dist}\left(\bar{x}, f^{-1}(\bar{y})\right) \leq \tilde{\kappa}\|\bar{y} - f(\bar{x})\|. \tag{3.4}$$

If $\bar{y} = f(\bar{x})$ the proof of the convexity of $f(S)$ is complete, because $\bar{x} \in S$. Otherwise, it happens that $\|\bar{y} - f(\bar{x})\| > 0$, so the inequality (3.4) entails the existence of $\hat{x} \in f^{-1}(\bar{y})$ such that

$$\|\hat{x} - \bar{x}\| < \kappa\|\bar{y} - f(\bar{x})\|.$$

By taking account of the estimate (2.2) in Remark 2.1 (i), as it is $[x_1, x_2] \in \mathrm{B}(x_0, \rho) \subseteq \text{int}\, \mathrm{B}(x_0, r_0)$, one consequently obtains

$$\|\hat{x} - \bar{x}\| < \kappa\frac{\ell}{8}\|x_1 - x_2\|^2,$$

that is $\hat{x} \in \mathrm{B}\left(\bar{x}, \frac{\kappa\ell}{8}\|x_1 - x_2\|^2\right)$. Since $S \in \mathcal{UC}_c^2(\mathbb{X})$ and the inequality (3.1) is in force, in the light of what observed in Remark 2.4 (iii) it follows

$$\mathrm{B}\left(\bar{x}, \frac{\kappa\ell}{8}\|x_1 - x_2\|^2\right) \subseteq S,$$

with the consequence that $\hat{x} \in S$ and hence $\bar{y} = f(\hat{x})$ turns out to belong to $f(S)$.

Second part: Let us prove now the assertion in the thesis. According to what noted in Remark 2.1 (ii), under the above hypotheses $f(S)$ is bounded. Fix $\epsilon \in (0, \text{diam}\, f(S))$ and take arbitrary y_1, $y_2 \in f(S)$, with $\|y_1 - y_2\| = \epsilon$. Let $\bar{y}$, x_1, x_2, $\bar{x}$ and $\hat{x}$ be as in the first part of the proof (it may happen that $\hat{x} = \bar{x}$). In order to prove that $f(S) \in \mathcal{UC}^2(\mathbb{Y})$, it is to be shown that, independently of y_1, $y_2 \in f(S)$ and ϵ, there exists $\gamma > 0$ such that $\mathrm{B}\left(\bar{y}, \gamma\epsilon^2\right) \subseteq f(S)$. Again recalling Remark 2.1 (ii), it is possible to define the positive real value

$$\beta = \sup_{x \in S} \|\mathrm{D}f(x)\|_{\mathcal{L}} + 1 < +\infty.$$

By virtue of inequality (3.1), it is possible to pick $\eta \in (0, c - \frac{\kappa\ell}{8})$ in such a way that

$$\hat{x} \in \mathrm{B}\left(\bar{x}, \frac{\kappa\ell}{8}\|x_1 - x_2\|^2\right) \subseteq \mathrm{B}\left(\bar{x}, \left(\frac{\kappa\ell}{8} + \eta\right)\|x_1 - x_2\|^2\right) \subseteq S.$$

From the last chain of inclusions, it readily follows that

$$\mathrm{B}\left(\hat{x}, \eta\|x_1 - x_2\|^2\right) \subseteq S.$$

Since, by the mean-value theorem, it is

$$\|y_1 - y_2\| \leq \beta\|x_1 - x_2\|,$$

one obtains

$$\epsilon^2 = \|y_1 - y_2\|^2 \leq \beta^2\|x_1 - x_2\|^2,$$

and hence $\mathrm{B}\left(\hat{x}, \eta\epsilon^2/\beta^2\right) \subseteq S$. Now, recall that f is open at a linear rate around x_0. Accordingly, as $S \subseteq \mathrm{B}(x_0, \rho)$, up to a further reduction in the value of $\eta > 0$ in such a way that $\eta\text{diam}^2 f(S)/\beta^2 < \rho$, one finds

$$\mathrm{B}\left(\bar{y}, \sigma\eta\frac{\epsilon^2}{\beta^2}\right) \subseteq f\left(\mathrm{B}\left(\hat{x}, \eta\frac{\epsilon^2}{\beta^2}\right)\right) \subseteq f(S)$$

(remember the inclusion (3.3)). Thus, since by construction σ, η and β are independent of y_1, y_2 and ϵ, one can conclude that

$$\delta_{f(S)}(\epsilon) \geq \frac{\sigma\eta}{\beta^2}\epsilon^2.$$

By arbitrariness of $\epsilon \in (0, \operatorname{diam} f(S))$, this completes the proof. □

A first comment to Theorem 3.2 concerns its hypothesis (iii), which seems to find no counterpart in the convexity principle due to B.T. Polyak (see [22, Theorem 2.1]). Such hypothesis postulates a uniform convexity property of S, which must be quantitatively adequate to the metric regularity of f and to the Lipschitz continuity of $\mathrm{D}f$ around x_0. Matching this condition is guaranteed for strongly convex sets (in particular, for balls) with a sufficiently small radius, provided that the underlying Banach space fulfils a certain uniform convexity assumption. This fact is clarified by the following

Corollary 3.3. *Let $f : \Omega \longrightarrow \mathbb{Y}$ be a mapping between Banach spaces, with Ω open nonempty subset of $\mathbb{X}$. Let $x_0 \in \Omega$ be such that:*
(i) *$(\mathbb{X}, \|\cdot\|)$ admits a modulus of convexity of power type 2;*
(ii) *$f \in \mathrm{C}^{1,1}(\operatorname{int} \mathrm{B}(x_0, r_0))$, for some $r_0 > 0$;*
(iii) *$\mathrm{D}f(x_0)$ is onto.*

Then, there exists $\rho \in (0, r_0)$ such that, for every r-convex set S, with $r \in [0, \rho)$ and $f(S)$ closed, it holds $f(S) \in \mathcal{UC}^2(\mathbb{Y})$.

Proof. By virtue of the hypothesis (i), according to Example 2.3 (ii), any r-convex set S belongs to $\mathcal{UC}^2(\mathbb{X})$, for every $r > 0$. More precisely, on account of the inequality (2.3), one has

$$\delta_S(\epsilon) \geq r\delta_{\mathbb{X}}\left(\frac{\epsilon}{r}\right) \geq \frac{\gamma}{r}\epsilon^2, \quad \forall \epsilon \in (0, 2r],$$

for some $\gamma > 0$. Therefore, in order for the hypothesis (iii) of Theorem 3.2 to be satisfied, it suffices to take

$$r < \frac{8\gamma}{\operatorname{reg}(f, x_0) \cdot \operatorname{lip}(\mathrm{D}f, \operatorname{int} \mathrm{B}(x_0, r_0)) + 1}.$$

Then, the thesis follows from Theorem 3.2. □

On the other hand, notice that Theorem 3.2 does not make any direct assumption on the Banach space $(\mathbb{X}, \|\cdot\|)$ (nonetheless, take into account what remarked at the end of Example 2.3 (i)). Furthermore, since any ball $\mathrm{B}(x_0, r)$ is a r-convex sets, it should be clear that Corollary 3.3 allows one to embed in the current theory the Polyak convexity principle and its refinement [26, Theorem 3.2].

Another comment to Theorem 3.2 deals with the topological assumption on the image $f(S)$. Of course, whenever $\mathbb{X}$ is a finite-dimensional Euclidean space, $f(S)$ is automatically closed, because S is compact and f is continuous on S. In an infinite-dimensional setting, the same issue becomes subtler. The closedness assumption thus appears also in the formulation of other results for the convexity of images of mappings between infinite-dimensional spaces (see [1, Theorem 2.2]). It is clear that, whenever $\mathrm{D}f(x_0)$ not only is onto but, in particular, is invertible, f turns out to be a diffeomorphism around x_0. As a consequence, for a proper $r_0 > 0$, any closed set $S \subseteq \mathrm{B}(x_0, r_0)$ has a closed image. Nevertheless, in the general setting of Theorem 3.2, to the best of the author's knowledge, the question of formulating sufficient conditions on f in order for $f(S)$ to be closed is still open. The next proposition, which is far removed from providing a solution to such a question, translates the topological assumption on the image $f(S)$ into variational terms.

Proposition 3.4. *Let $f : \Omega \longrightarrow \mathbb{Y}$ be a mapping between Banach spaces, with Ω open nonempty subset of $\mathbb{X}$, and let $x_0 \in \Omega$. Suppose that:*

(i) *f is continuous in $\mathrm{B}(x_0, r_0)$, for some $r_0 > 0$;*
(ii) *the function $x \mapsto \mathrm{dist}\left(x, f^{-1}(y)\right)$ is weakly lower semicontinuous, for every $y \in \mathrm{B}(f(x_0), r_0)$;*
(iii) *$(\mathbb{X}, \|\cdot\|)$ is reflexive;*
(iv) *f is metrically regular around x_0.*

Then, there exists $\rho \in (0, r_0)$ such that, for every closed convex set $S \subseteq \mathrm{B}(x_0, \rho)$, $f(S)$ is closed.

Proof. Since by the hypothesis (iv) f is metrically regular around x_0, there exist positive real $r \in (0, r_0)$ and κ such that

$$\mathrm{dist}\left(x, f^{-1}(y)\right) \leq \kappa\|f(x) - y\|, \quad \forall x \in \mathrm{B}(x_0, r), \ \forall y \in \mathrm{B}(f(x_0), r). \tag{3.5}$$

By the continuity of f at x_0, there exists $\rho \in (0, r)$ such that

$$f(x) \in \mathrm{B}(f(x_0), r), \quad \forall x \in \mathrm{B}(x_0, \rho).$$

Thus, whenever $S \subseteq \mathrm{B}(x_0, \rho)$, one has $f(S) \subseteq \mathrm{B}(f(x_0), r)$.

Now, suppose that $S \subseteq \mathrm{B}(x_0, \rho)$ is a closed convex set and take an arbitrary $y \in \mathrm{cl}\, f(S) \subseteq \mathrm{B}(f(x_0), r)$. Let $(y_n)_n$ be a sequence in $f(S)$, such that $y_n \longrightarrow y$ as $n \to \infty$. As $y_n \in f(S)$, there exists a sequence $(x_n)_n$ in S such that $y_n = f(x_n)$, for each $n \in \mathbb{N}$. Notice that, since $x_n \in S \subseteq \mathrm{B}(x_0, \rho) \subseteq \mathrm{B}(x_0, r)$ and $y \in \mathrm{cl}\, f(S) \subseteq \mathrm{B}(f(x_0), r)$, the inequality (3.5) applies, namely

$$\mathrm{dist}\left(x_n, f^{-1}(y)\right) \leq \kappa\|f(x_n) - y\| = \kappa\|y_n - y\|, \quad \forall n \in \mathbb{N}. \tag{3.6}$$

This shows that $\mathrm{dist}\left(x_n, f^{-1}(y)\right) \longrightarrow 0$ as $n \to \infty$ and therefore

$$\inf_{x \in S} \mathrm{dist}\left(x, f^{-1}(y)\right) = 0.$$

As a closed convex set, S is also weakly closed. Moreover, as a bounded subset of a reflexive Banach space, S is weakly compact. Thus, since $y \in \mathrm{B}(f(x_0), r_0)$, by virtue of the hypothesis (ii), there must exist $\tilde{x} \in S$ such that

$$\mathrm{dist}\left(\tilde{x}, f^{-1}(y)\right) = 0.$$

Since f is continuous, the last inequality entails that $\tilde{x} \in f^{-1}(y)$. This leads to conclude that $y \in f(S)$, thereby completing the proof. □

The hypothesis (ii) in Proposition 3.4 happens to be always satisfied if f is a linear mapping. In the nonlinear case, the situation is expected to be much more complicate.

Let $C \subseteq \mathbb{Y}$ be a closed convex cone with apex at $\mathbf{0}$ and let $S \subseteq \mathbb{X}$ be nonempty and convex. Recall that a mapping $f : S \longrightarrow \mathbb{Y}$ is said to be *convex-like* on S with respect to C if for every $x_1,\, x_2 \in S$ and $t \in [0, 1]$, there exists $x_t \in S$ such that

$$(1 - t)f(x_1) + tf(x_2) \in f(x_t) + C.$$

Convex-likeness is a generalization of the notion of C-convexity of mappings taking values in partially ordered vector spaces. It should be evident that, when $\mathbb{Y} = \mathbb{R}$, $C = [0, +\infty)$ and $x_t = (1-t)x_1 + tx_2$, the above inclusion reduces to the well-known inequality defining the convexity of a functional. The class of convex-like mappings has found a large employment in optimization and related topics. For instance, if $\mathbb{R}^m$ and $C = \mathbb{R}^m_+$ it is readily seen that this class includes all mappings $f = (f_1, \dots, f_m)$, having each component $f_i : S \longrightarrow \mathbb{R}$, $i = 1, \dots, m$ convex on a convex set. For a detailed discussion about the notion of convex-likeness of mappings, its variants and their impact on the study of variational problems, the reader can refer to [15]. The next corollary, which can be achieved as a direct consequence of Theorem 3.2, reveals that any $\mathrm{C}^{1,1}$ smooth mapping behave as a convex-like mapping on uniformly convex sets of class $\mathcal{UC}^2_c(\mathbb{X})$ near a regular point.

Corollary 3.5. *Let $f : \Omega \longrightarrow \mathbb{Y}$ be a mapping between Banach spaces, $x_0 \in \Omega$ and $c > 0$. If f, x_0 and c satisfy all hypotheses of Theorem 3.2, then there exists $\rho > 0$ such that, for every $S \in \mathcal{UC}^2_c(\mathbb{X})$, with $S \subseteq \mathrm{B}\,(x_0, \rho)$ and $f(S)$ closed, and every cone $C \subseteq \mathbb{Y}$, the mapping $f : S \longrightarrow \mathbb{Y}$ is convex-like on S with respect to C.*

Proof. The thesis follows at once by Theorem 3.2, from being

$$(1 - t)f(x_1) + tf(x_2) \in f(S) \subseteq f(S) + C, \quad \forall x_1,\, x_2 \in S,\ \forall t \in [0, 1].$$

□

4. Applications to optimization

Throughout this section, applications of Theorem 3.2 will be considered to the study of constrained optimization problems, having the following format

$$(\mathcal{P}) \qquad \min_{x\in S} \varphi(x) \quad \text{subject to} \quad g(x) \in C,$$

where $\varphi : \mathbb{X} \longrightarrow \mathbb{R}$ and $g : \mathbb{X} \longrightarrow \mathbb{Y}$ are given functions between Banach spaces, $S \subseteq \mathbb{X}$ and $C \subseteq \mathbb{Y}$ are given (nonempty) closed and convex sets. Such a format is frequently employed in the literature for subsuming under a general treatment a broad spectrum of finite and infinite-dimensional extremum problems, with various kinds of constraints. The feasible region of problem $(\mathcal{P})$ will be henceforth denoted by R, i.e. $R = S \cap g^{-1}(C)$.

According to a long-standing approach in optimization, now recognized as ISA (acronym standing for Image Space Analysis), the analysis of several issues related to problem $(\mathcal{P})$ can be performed by associating with $(\mathcal{P})$ and with an element $x_0 \in R$ the mapping $f_{\mathcal{P},x_0} : \mathbb{X} \longrightarrow \mathbb{R} \times \mathbb{Y}$, which is defined by

$$f_{\mathcal{P},x_0}(x) = (\varphi(x) - \varphi(x_0), g(x))$$

(see, for instance, [13] and references therein). It is natural to believe that the mapping $f_{\mathcal{P},x_0}$ inherits certain structural features of the given problem. Such issues as the solution existence, optimality conditions, duality, and so on, can be investigated by studying relationships between the two subsets of the space $\mathbb{R}\times\mathbb{Y}$, namely $f_{\mathcal{P},x_0}(S)$ and $Q=(-\infty,0)\times C$, associated with $(\mathcal{P})$.

Remark 4.1. Directly from the above constructions, it is possible to prove the following well-known facts:
(i) $x_0\in R$ is a global solution to $(\mathcal{P})$ iff $f_{\mathcal{P},x_0}(S)\cap Q=\varnothing$;
(ii) $x_0\in R$ is a local solution to $(\mathcal{P})$ iff there exists $r>0$ such that $f_{\mathcal{P},x_0}(S\cap \mathrm{B}(x_0,r))\cap Q=\varnothing$.

The above facts have been largely employed as a starting point for formulating optimality conditions within ISA. Another relevant property connected with optimality is openness at a linear rate. Its presence, indeed, has been observed to be in contrast with optimality (see, for instance, the so-called noncovering principle in [14]). Below, a lemma related to this phenomenon, which will be exploited in the proof of the next result, is presented in full detail.

Lemma 4.2. *With reference to a problem $(\mathcal{P})$, suppose that the mapping $f_{\mathcal{P},x_0}$ is open at a linear rate around $x_0\in R$ and $x_0\in\operatorname{int}S$. Then, x_0 is not a local solution to $(\mathcal{P})$.*

Proof. By the hypothesis, according to Definition 2.6 there exist positive constants δ, ζ, and σ such that, if taking in particular $x=x_0$ in inclusion (2.6), it holds

$$f_{\mathcal{P},x_0}(\mathrm{B}(x_0,r))\supseteq \mathrm{B}(f_{\mathcal{P},x_0}(x_0),\sigma r)\cap \mathrm{B}(f_{\mathcal{P},x_0}(x_0),\zeta), \quad \forall r\in[0,\delta].$$

Notice that, if $r<\zeta/\sigma$, then the above inclusion reduces to

$$f_{\mathcal{P},x_0}(\mathrm{B}(x_0,r))\supseteq \mathrm{B}(f_{\mathcal{P},x_0}(x_0),\sigma r)=\mathrm{B}((0,g(x_0)),\sigma r). \tag{4.1}$$

Since $x_0\in\operatorname{int}S$, there exists $r_0>0$ such that $\mathrm{B}(x_0,r_0)\subseteq S$. Now, fix an arbitrary $r\in(0,\min\{r_0,\zeta/\sigma\})$ and pick $t\in(0,\sigma r)$. Then, on the account of inclusion (4.1), there exists $x_r\in\mathrm{B}(x_0,r)$ such that

$$f_{\mathcal{P},x_0}(x_r)=(-t,g(x_0))\in\mathrm{B}((0,g(x_0)),\sigma r),$$

that is

$$\varphi(x_r)-\varphi(x_0)=-t<0 \qquad \text{and} \qquad g(x_r)=g(x_0)\in C.$$

This means that $x_r\in S\cap g^{-1}(C)$ and $\varphi(x_r)<\varphi(x_0)$, what contradicts the local optimality of x_0 for $(\mathcal{P})$, by arbitrariness of r. □

The next theorem, which extends a similar result established in [26, Theorem 3.2], provides an answer to the question of solution existence for problem $(\mathcal{P})$ and, at the same time, furnishes an optimality condition for detecting a solution. In order to formulate such a theorem, let us denote by $\mathrm{N}(C,\bar{y})=\{y^*\in\mathbb{Y}^*:\ \langle y^*,y-\bar{y}\rangle\le 0,\quad \forall y\in C\}$ the normal cone to C at $\bar{y}$ in the sense of convex analysis. Besides, let us denote by $\mathrm{L}:\mathbb{Y}^*\times\mathbb{X}\longrightarrow\mathbb{R}$ the Lagrangian function associated with problem $(\mathcal{P})$, i.e.

$$\mathrm{L}(y^*,x)=\varphi(x)+\langle y^*,g(x)\rangle.$$

The proof, whose main part is given for the sake of completeness, adapts an argument already exploited in [26]. It derives solution existence from the weak compactness of the problem image and the optimality condition by a linear separation technique. In both the cases, convexity is the geometrical property that makes this possible.

Theorem 4.3. *Given a problem* $(\mathcal{P})$, *let* $x_0 \in g^{-1}(C)$ *and let* c *be a positive real. Suppose that:*

(i) $(\mathbb{Y}, \|\cdot\|)$ *is a reflexive Banach space;*
(ii) $\varphi,\, g \in \mathrm{C}^{1,1}(\operatorname{int} \mathrm{B}\,(x_0, r_0))$, *for some* $r_0 > 0$ *and* $\mathrm{D}f_{\mathcal{P},x_0}(x_0)$ *is onto;*
(iii) *it holds*

$$\frac{\operatorname{reg}(f_{\mathcal{P},x_0}, x_0) \cdot \operatorname{lip}(\mathrm{D}f_{\mathcal{P},x_0}, \operatorname{int} \mathrm{B}\,(x_0, r_0))}{8} < c. \tag{4.2}$$

Then, there exists $\rho \in (0, r_0)$ *such that, for every* $S \in \mathcal{UC}^2_c(\mathbb{X})$, *with* $x_0 \in \operatorname{int} S \subseteq \mathrm{B}\,(x_0, \rho)$ *and* $f_{\mathcal{P},x_0}(S)$ *closed, one has*

(t) *there exists a global solution* $\bar{x}_S \in R$ *to* $(\mathcal{P})$;
(tt) $\bar{x}_S \in \operatorname{bd} S$ *and hence* $\bar{x}_S \in \operatorname{bd} R$;
(ttt) *there exists* $y^*_S \in \mathrm{N}(C, g(\bar{x}_S))$ *such that*

$$\mathrm{L}(y^*_S, \bar{x}_S) = \min_{x \in S} \mathrm{L}(y^*_S, x).$$

Proof. (t) Under the hypotheses (ii) and (iii), one can apply Theorem 3.2. If $\rho > 0$ is as in the thesis Theorem 3.2, fix a set $S \in \mathcal{UC}^2_c(\mathbb{X})$ satisfying all requirements in the above statement. Then its image $f_{\mathcal{P},x_0}(S)$ turns out to be a convex, closed and bounded subset of $\mathbb{R} \times \mathbb{Y}$, with nonempty interior. The existence of a global solution to $(\mathcal{P})$ will be achieved by proving that an associated minimization problem in the space $\mathbb{R} \times \mathbb{Y}$ does admit a global solution. To do so, define

$$\tau = \inf\{t :\ (t, y) \in f_{\mathcal{P},x_0}(S) \cap Q\}.$$

Notice that $x_0 \in R$. Since $\mathrm{D}f_{\mathcal{P},x_0}(x_0)$ is onto, by the Lyusternik-Graves theorem the mapping $f_{\mathcal{P},x_0}$ too is open at a linear rate around x_0. Thus, since $x_0 \in \operatorname{int} S$, in the light of Lemma 4.2 x_0 must fail to be a local (and hence, a fortiori, global) solution to $(\mathcal{P})$. Consequently, according to what observed in Remark 4.1 (i), it must be

$$f_{\mathcal{P},x_0}(S) \cap Q \neq \varnothing.$$

This implies that $\tau < +\infty$. Furthermore, if setting

$$\bar{\tau} = \inf\{t :\ (t, y) \in f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q\}, \tag{4.3}$$

it is possible to see that actually it is $\bar{\tau} = \tau$. Indeed, since x_0 is not a solution to $(\mathcal{P})$, there exists $\hat{x} \in R$ such that $\varphi(\hat{x}) - \varphi(x_0) < 0$, and so $f_{\mathcal{P},x_0}(\hat{x}) = (\varphi(\hat{x}) - \varphi(x_0), g(\hat{x})) \in f_{\mathcal{P},x_0}(S) \cap Q$. As $f_{\mathcal{P},x_0}(S) \cap Q \subseteq f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q$, it follows that $\bar{\tau} \leq \tau \leq \varphi(\hat{x}) - \varphi(x_0) < 0$. Hence, for any $\epsilon \in (0, -\bar{\tau})$ there exists $(t_\epsilon, y_\epsilon) \in f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q$ such that $t_\epsilon < \bar{\tau} + \epsilon < 0$. Noting that $\operatorname{cl} Q = (-\infty, 0] \times C$, this implies that $(t_\epsilon, y_\epsilon) \in f_{\mathcal{P},x_0}(S) \cap Q$ and consequently that $\bar{\tau} \leq \tau \leq t_\epsilon < \bar{\tau} + \epsilon < 0$. Letting $\epsilon \to 0^+$, one obtains $\bar{\tau} = \tau$.

Now, as the set $f_{\mathcal{P},x_0}(S)$ is closed, convex and bounded, so is its subset $f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q$. The boundedness of the latter implies that $\bar{\tau} > -\infty$. Moreover, by virtue of the hypothesis (i), $f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q$ turns out to be weakly compact. Since the projection mapping $\Pi_{\mathbb{R}} : \mathbb{R} \times \mathbb{Y} \longrightarrow \mathbb{R}$, given by $\Pi_{\mathbb{R}}(t, y) = t$ is continuous and convex, it is also weakly l.s.c., with the consequence that the infimum defined in (4.3) is actually attained at some $(\bar{t}, \bar{y}) \in f_{\mathcal{P},x_0}(S) \cap \operatorname{cl} Q$. This means that there exists $\bar{x}_S \in S$ such that

$$\tau = \bar{\tau} = \bar{t} = \varphi(\bar{x}_S) - \varphi(x_0) \quad \text{ and } \bar{y} = g(\bar{x}_S) \in C.$$

Let us show that $\bar{x}_S$ is a global solution to $(\mathcal{P})$. Assume to the contrary that there is $\hat{x} \in R$ such that $\varphi(\hat{x}) < \varphi(\bar{x}_S)$. Then, one finds

$$\begin{aligned} \hat{t} &= \varphi(\hat{x}) - \varphi(x_0) = \varphi(\hat{x}) - \varphi(\bar{x}_S) + \varphi(\bar{x}_S) - \varphi(x_0) \\ &< \varphi(\bar{x}_S) - \varphi(x_0) = \bar{t} = \bar{\tau} = \tau. \end{aligned}$$

Since it is $\hat{x} \in R$, then $\hat{x} \in S$ and $\hat{y} = g(\hat{x}) \in C$, wherefrom one has $(\hat{t}, \hat{y}) \in f_{\mathcal{P},x_0}(S) \cap Q$, which contradicts the definition of τ.

(tt) To prove that $\bar{x}_S$ belongs to $\operatorname{bd} S$, notice that $(\bar{t}, \bar{y}) = f_{\mathcal{P},x_0}(\bar{x}_S) \in \operatorname{bd} f_{\mathcal{P},x_0}(S)$. Then, by recalling what mentioned in Remark 2.7 (ii), this assertion follows from the openness at a linear rate of $f_{\mathcal{P},x_0}$ around x_0.

(ttt) Again remembering Remark 4.1 (i), by the global optimality of $\bar{x}_S$, it results in

$$f_{\mathcal{P},\bar{x}_S}(S) \cap Q = \varnothing. \tag{4.4}$$

As one readily checks, it holds

$$f_{\mathcal{P},\bar{x}_S}(S) = f_{\mathcal{P},x_0}(S) + (\varphi(x_0) - \varphi(\bar{x}_S), \mathbf{0}),$$

that is to say $f_{\mathcal{P},\bar{x}_S}(S)$ is a translation of $f_{\mathcal{P},x_0}(S)$. Therefore, $f_{\mathcal{P},\bar{x}_S}(S)$ too is a closed, bounded, convex subset of $\mathbb{R} \times \mathbb{Y}$, with nonempty interior. Since (4.4) is true, the Eidelheit theorem makes it possible to linearly separate $f_{\mathcal{P},\bar{x}_S}(S)$ and $\operatorname{cl} Q$. In other terms, this means the existence of a pair $(\gamma, y^*) \in (\mathbb{R} \times \mathbb{Y}) \backslash \{(0, \mathbf{0}^*)\}$ and $\alpha \in \mathbb{R}$ such that

$$\gamma(\varphi(x) - \varphi(\bar{x}_S)) + \langle y^*, g(x) \rangle \geq \alpha, \quad \forall x \in S,$$

and

$$\gamma t + \langle y^*, y \rangle \leq \alpha, \quad \forall (t, y) \in \operatorname{cl} Q = (-\infty, 0] \times C.$$

The rest of the proof relies on a standard usage of the last inequalities and does not need to devise any specific adaptation. □

Theorem 4.3 describes the local behaviour of a nonlinear optimization problem $(\mathcal{P})$ near a point $x_0 \in (\operatorname{int} S) \cap g^{-1}(C)$, around which the condition (4.2) linking the modulus of convexity of S, the regularity behaviour of $f_{\mathcal{P},x_0}$ and the Lipschitz continuity of its derivative happens to be satisfied: $(\mathcal{P})$ admits a global solution, which lies at the boundary of the feasible region and can be detected by minimizing the Lagrangian function. The reader should notice that globality of a solution and its characterization as a minimizer of a the Lagrangian function are phenomena typically occurring in convex optimization. Instead, they generally fail to occur in nonlinear optimization, where optimality conditions are usually only necessary or sufficient, and frequently expressed in terms of Lagrangian stationary by means of first-order derivative.

Another typical phenomenon arising in convex optimization is the vanishing of the duality gap, i.e. the vanishing of the value

$$\operatorname{gap}(\mathcal{P}) = \inf_{x \in S} \sup_{y^* \in C^{\ominus}} \mathrm{L}(y^*, x) - \sup_{y^* \in C^{\ominus}} \inf_{x \in S} \mathrm{L}(y^*, x),$$

where $C^{\ominus} = \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle \leq 0\}$ is the dual cone to C. Such a circumstance, which can be proved to take place in convex programming under proper qualification conditions, is known as strong (Lagrangian) duality. In the current setting, it can be readily achieved as a consequence of Theorem 4.3, without the need of extra assumptions, apart from the cone structure now imposed on the set C.

Corollary 4.4. *Given a problem* $(\mathcal{P})$*, suppose that* C *is a closed convex cone. Under the hypothesis of Theorem 4.3, it holds*

$$\operatorname{gap}(\mathcal{P}) = 0$$

and there exists a pair $(y^*_S, \bar{x}_S) \in C^{\ominus} \times R$*, which is a saddle point of* L*, i.e.*

$$\mathrm{L}(y^*, \bar{x}_S) \leq \mathrm{L}(y^*_S, \bar{x}_S) \leq \mathrm{L}(y^*_S, x), \quad \forall (y^*, x) \in C^{\ominus} \times S.$$

Proof. Let $\bar{x}_S$ and y^*_S be as in the thesis of Theorem 4.3. Since C is a closed convex cone, $2g(\bar{x}_S)$ and $\mathbf{0}$ belong to C. By recalling that $y^*_S \in \mathrm{N}(C, g(\bar{x}_S))$, one has

$$\langle y^*_S, y - g(\bar{x}_S)\rangle \leq 0, \quad \forall y \in C.$$

By replacing y with $2g(\bar{x}_S)$ and $\mathbf{0}$ in last inequality, one easily shows that $\langle y^*_S, g(\bar{x}_S)\rangle = 0$ and hence $y^*_S \in C^{\ominus}$. The rest of the thesis then follows at once. □

The above applications of Theorem 3.2 demonstrate that, even in the absence of convexity assumptions on the functional data of problem $(\mathcal{P})$, some good phenomena connected with convexity may still appear.

Example 4.5. With reference to the problem format $(\mathcal{P})$, let $\mathbb{X} = \mathbb{R}^2$, $\mathbb{Y} = \mathbb{R}$, $C = \{0\}$, and let $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined respectively by

$$\varphi(x) = x_1^2 - x_2^2, \qquad g(x) = x_1^2 + x_2^2 - 1.$$

Take $x_0 = (1/\sqrt{2}, 1/\sqrt{2}) \in g^{-1}(0) = \mathbb{S}$ and $S = \mathrm{B}(x_0, r)$. With the above choice of data, the problem falls out of the realm of convex optimization: the objective function φ is evidently not convex as well as the feasible region $R = S \cap \mathbb{S}$, for every $r > 0$. Throughout the present example, $\mathbb{R}^2$ is supposed to be equipped with its Euclidean space structure, so that

$$\delta_{\mathbb{R}^2}(\epsilon) \geq \frac{\epsilon^2}{8}, \quad \forall \epsilon \in (0, 2].$$

Therefore, $S = \mathrm{B}(x_0, r) \in \mathcal{U}C^2(\mathbb{R}^2)$ and, according to the estimate in (2.3), one finds

$$\delta_{\mathrm{B}(x_0,r)}(\epsilon) \geq r\delta_{\mathbb{R}^2}\left(\frac{\epsilon}{r}\right) = \frac{\epsilon^2}{8r},$$

that is $\mathrm{B}(x_0, r) \in \mathcal{U}C^2_{1/8r}(\mathbb{R}^2)$, for every $r > 0$. Clearly, the function $f_{\mathcal{P},x_0} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, which is given in this case by

$$f_{\mathcal{P},x_0}(x) = \begin{pmatrix} x_1^2 - x_2^2 \\ x_1^2 + x_2^2 - 1 \end{pmatrix},$$

satisfies the smoothness hypothesis of Theorem 4.3. In particular, since it is

$$\mathrm{D}f_{\mathcal{P},x_0}(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_1 & x_2 \end{pmatrix},$$

it results in

$$\operatorname{reg}(f_{\mathcal{P},x_0}, x_0) = \|\mathrm{D}f_{\mathcal{P},x_0}(x_0)^{-1}\|_{\mathcal{L}} = \left\| \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \right\|_{\mathcal{L}} = \frac{1}{2}.$$

On the other hand, since the mapping $\mathrm{D}f_{\mathcal{P},x_0} : \mathbb{R}^2 \longrightarrow \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ is linear in this case, one finds

$$\begin{aligned} \operatorname{lip}(\mathrm{D}f_{\mathcal{P},x_0}, \mathbb{R}^2) &= \|\mathrm{D}f_{\mathcal{P},x_0}\|_{\mathcal{L}} = \max_{u \in \mathbb{S}} \|\mathrm{D}f_{\mathcal{P},x_0}(u)\|_{\mathcal{L}} \\ &= \max_{u \in \mathbb{S}} \max_{v \in \mathbb{S}} \|\mathrm{D}f_{\mathcal{P},x_0}(u)v\| = 2\sqrt{2}. \end{aligned}$$

Consequently, the condition (4.2) becomes

$$\frac{\frac{1}{2} \cdot 2\sqrt{2}}{8} < \frac{1}{8r}.$$

Thus, for every $r < 1/\sqrt{2}$, by virtue of Theorem 4.3 assertions (t) – (ttt) hold. In particular, it is not difficult to check (for instance, by means of a level set inspection) that for every $S = \mathrm{B}\,(x_0, r)$, with $r < 1/\sqrt{2}$, the unique (global) solution $\bar{x}_S$ of the related problem lies in $\operatorname{bd} S$. Notice that this fails to be true if $r > \sqrt{2 - \sqrt{2}} = \|(0,1) - x_0\| > 1/\sqrt{2}$, in which case the solution $\bar{x}_S = (0,1)$ belongs to $\operatorname{int} \mathrm{B}\,(x_0, r) = \operatorname{int} S$.

Acknowledgements

The author thanks an anonymous referee for valuable remarks, which help him to considerably improve the quality of his paper; in particular, for pointing out a crucial gap in the original proof of Theorem 4.3 as well as several inaccuracies.

References

[1] A. Baccari and B. Samet, *An extension of Polyak's theorem in a Hilbert space*, J. Optim. Theory Appl. **140** (2009), 409–418.

[2] M. V. Balashov and D. Repovš, *Uniform convexity and the splitting problem for selections*, J. Math. Anal. Appl. **360** (2009), 307–316.

[3] M. V. Balashov and D. Repovš, *Uniformly convex subsets of the Hilbert space with modulus of convexity of the second order*, J. Math. Anal. Appl. **377** (2011), 754–761.

[4] N. A. Bobylev, S.V. Emel'yanov and S.K. Korovin, *Convexity of images of convex sets under smooth maps. Nonlinear dynamics and control. 2*, Comput. Math. Model. **15** (2004), 213–222.

[5] J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (1936), 396–414.

[6] J. Diestel, *Geometry of Banach Spaces - Selected Topics*, Springer-Verlag, Berlin-New York, 1975.

[7] L. L. Dines, *On the mapping of quadratic forms*, Bull. Amer. Math. Soc. **47** (1941), 494–498.

[8] 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.

[9] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, New York, 2014. and

[10] A. Dymarsky, *Convexity of a small ball under quadratic map*, Linear Algebra Appl. **488** (2016), 109–123.

[11] P. Enflo, *Banach spaces which can be given an equivalent uniformly convex norm*, Israel J. Math. **13** (1972), 281–288.

[12] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant and V. Zizler, *Functional Analysis and Infinite-Dimensional Geometry*, Springer-Verlag, New York, 2001.

[13] F. Giannessi, *Constrained Optimization and Image Space Analysis. Vol. 1. Separation of Sets and Optimality Conditions*, Springer, New York, 2005.

[14] A. D. Ioffe, *Metric regularity – a survey Part 1. Theory*, J. Aust. Math. Soc. **101** (2016), 188–243.

[15] G. Mastroeni and T. Rapcsák, *On convex generalized systems*, J. Optim. Theory Appl. **104** (2000), 605–627.

[16] V. D. Milman, *Geometric theory of Banach spaces. II. Geometry of the unit ball*, Uspekhi Mat. Nauk **26** (1971), 73–149.

[17] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation. I. Basic Theory*, Springer-Verlag, Berlin, 2006.

[18] E. S. Polovinkin, *On properties of strongly convex sets*, Modeling of Control and Data Processing Processes, Mosk. Fiz.-Tekh. Inst. (1994), 182–189.

[19] E. S. Polovinkin, *Strongly convex analysis*, Sb. Math. **187** (1996), 259–286.

[20] B. T. Polyak, *Existence theorems and convergence of minimizing sequences in extremum problems with restrictions*, Soviet Math. **7** (1966), 72–75.

[21] B. T. Polyak, *Convexity of quadratic transformations and its use in control and optimization*, J. Optim. Theory Appl. **99** (1998), , 553–583.

[22] B. T. Polyak, *Convexity of Nonlinear Image of a Small Ball with Applications to Optimization*, Set-Valued Anal. **9** (2001), 159–168.

[23] B. T. Polyak, *Local programming*, Zh. Vychisl. Mat. Mat. Fiz. **41** (2001), 1259–1266.

[24] B. T. Polyak, *The convexity principle and its applications*, Bull. Braz. Math. Soc. (N.S.) **34** (2003), 59–75.

[25] G. Raĭsig, *Convexity of the reachable sets of control systems*, Autom. Remote Control **68** (2007), 1527–1543.

[26] A. Uderzo, *On the Polyak convexity principle and its application to variational analysis*, Nonlinear Anal. **91** (2013), 60–71.

[27] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

Manuscript received November 24 2017
revised March 29 2018

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