ON A CLASS OF CONVEX SETS WITH CONVEX IMAGES AND ITS APPLICATION TO NONCONVEX OPTIMIZATION

A. UDERZO

Abstract. In the present paper, conditions under which the images of uniformly convex sets through $C^{1,1}$ regular mappings between Banach spaces remain convex are established. These conditions are expressed by a certain quantitative relation between the modulus of convexity of a given set and the global regularity behaviour of the mapping on it. Such a result enables one to extend to a wide subclass of convex sets the Polyak’s convexity principle, which was originally concerned with images of small balls around points of Hilbert spaces. In particular, the crucial phenomenon of the preservation of convexity under regular $C^{1,1}$ transformations is shown to include the class of $r$-convex sets, where the value of $r$ depends on the regularity behaviour of the involved transformation. Two consequences related to nonconvex optimization are discussed: the first one is a sufficient condition for the global solution existence for infinite-dimensional constrained extremum problems; the second one provides a zero-order Lagrangian type characterization of optimality in nonlinear mathematical programming.

1. Introduction

In mathematical programming, control theory and in other areas of applied mathematics there is a tremendous need of convexity. In particular, convexity for sets is a geometric property without which entire approaches to important problems would be not conceivable, as well as a number of solution methods would be not practicable. This explains the active interest uninterruptedly maintained on that notion, which has been thereby subject to intensive investigations, from different viewpoints. Recognizing the irremediable lack of convexity in various circumstances stimulated researchers to develop alternative tools of analysis and original theories (nonconvex subdifferential/normal calculi, nonlinear separation techniques, extremal principles as a variational counterpart of linear separation, generalized convexity, like for instance in [7, 18, 29]), which in some cases contributed to deepen the knowledge about convexity itself.

Despite such an interest, not much is known so far about the preservation of convexity under transformations of sets. Whereas it is evident that linear mappings carry convex sets to convex sets, the question becomes subtler passing from the linear to the nonlinear case. In fact, simple counterexamples show that, already in the Euclidean plane $\mathbb{R}^2$, general convexity of sets fails to be preserved through nonlinear (even second degree polynomial) mappings. So one is led to seek for the preservation of convexity among classes of sets having stronger properties. A direction in which to move in such a search seems to be indicated by the Polyak’s convexity principle (see [24, 26]). This principle states that $C^{1,1}$ mappings between Hilbert spaces, which are regular at a given point, carry balls centered at the same point to convex sets, provided the radius of the balls is sufficiently small. A perusal of the proof of the Polyak’s principle reveals that rotundity of a set seems to be a crucial property in preserving its convexity under transformations. Instead, sets whose boundary contains line segments seem to be more exposed to the loss of convexity. To a deeper view, the mentioned convexity principle results from a proper interplay between rotundity of sets and regularity of smooth mappings.

A subclass of convex sets having good rotundity properties is that of uniformly convex sets. Roughly speaking, their rotundity reproduces that of the balls in uniformly convex Banach spaces. Curiously enough, such class of convex sets was introduced several decades ago by B.T. Polyak himself. Another class is that of strongly convex sets. Such sets are generated by intersection of balls with fixed radius. The aim of the present paper is to study conditions linking the rotundity property of sets in the aforementioned classes and...
the regularity of $C^{1,1}$ mappings, in such a way to achieve a nonlocal version of the Polyak’s convexity principle. A key concept in this study is a nonlocal form of metric regularity, called metric regularity on sets, and the related modulus. This concept refers to a notion, which provides a quantitative description of the surjectivity behaviour of a mapping on a whole set. It turns out that the constants that enable one to formulate sufficient conditions for metric regularity on sets play a certain role in analyzing the phenomenon of the convexity preservation. Based on these elements, the resulting conditions appear to be different from other nonlocal versions of the Polyak’s convexity principle, which has been obtained recently in [6] [27] [28].

The material presented in the paper is organized according to the following scheme. In Section 2 the notions of uniformly convex and strongly convex sets are introduced. Then properties which are relevant to the subsequent analysis are briefly recalled. A characterization for finite-dimensional uniformly convex sets is also obtained in terms of an extremality condition on their boundary. Section 3 is aimed at providing a sufficient condition for metric regularity on sets adequate to $C^{1,1}$ mappings between Banach spaces. This is obtained via a coincidence point result in metric spaces that allows one to extend the Milyutin theorem to the context under study. In Section 4 the main results of the paper, that is a nonlocal version of the Polyak’s convexity principle both for uniformly convex and for strongly convex sets, are established and discussed. Its specialization to quadratic transformations in Euclidean spaces is considered in detail. In Section 5 consequences of the main results are investigated with reference to nonconvex optimization problems. The aim of this kind of investigations is to try to restore in a nonlocal form some of the achievements of “local programming” (in the sense of [25]). By this term the theory is meant which emerges in connection with a special class of nonlinear optimization problems, exhibiting a local behaviour typical of convex optimization problems, even in the absence of convexity assumptions on their data.

2. Uniformly convex and strongly convex sets

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 $r \geq 0$, $B(x_0, r) = \{x \in X : d(x, x_0) \leq r\}$ 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 $B$, whereas the unit sphere by $S$. The distance of $x_0 \in X$ from a set $S \subseteq X$ is denoted by $\text{dist}(x_0, S)$. The diameter of a set $S \subseteq X$ is defined as $\text{diam} S = \sup\{d(x_1, x_2) : x_1, x_2 \in S\}$. By $\text{int} S$ and $\text{bd} S$ the topological interior and the boundary of a set $S$ are marked, respectively. If $S$ is a subset of a Banach space $(X, \| \cdot \|)$, $\text{ext} S$ denotes the set of all extreme points of $S$, in the sense of convex analysis; $0$ stands for the null element of $X$ and $[x_1, x_2]$ denotes the closed line segment with endpoints $x_1, x_2 \in X$. Given a function $h : X \to Y$ 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 $d(h(x_1), h(x_2)) \leq \ell d(x_1, x_2)$, for every $x_1, x_2 \in U$. The infimum over all values $\ell$ making the last inequality satisfied on $U$ is called modulus of Lipschitz continuity of $h$ on $U$ and is denoted by $\text{lip}(h, U)$. The Banach space of all bounded linear operators between the Banach spaces $X$ and $Y$, equipped with the operator norm, is denoted by $(\mathcal{L}(X, Y), \| \cdot \|_\mathcal{L})$. If, in particular, it is $Y = \mathbb{R}$, the simpler notation $(\mathcal{K}(X), \| \cdot \|_*)$ is used. Given a mapping $f : \Omega \to Y$, with $\Omega$ open subset of $X$, and $x_0 \in \Omega$, the Gateaux derivative of $f$ at $x_0$ is denoted by $Df(x_0)$. If $f$ is Gateaux differentiable at each point of $\Omega$ and mapping $Df : \Omega \to \mathcal{L}(X, Y)$ is Lipschitz continuous on $\Omega$, $f$ is said to be of class $C^{1,1}(\Omega)$.

Let $(X, \| \cdot \|)$ be a real Banach space. The main result of the paper relates to a property of persistence of convexity, which can be observed for a certain class of convex sets. This class was introduced in [22]. It can be defined as follows.

**Definition 2.1.** (i) Let $S \subseteq X$ be a nonempty, closed and convex subset of a real Banach space. The function $\delta_S : [0, \text{diam } S] \to [0, +\infty)$ defined by

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

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

(ii) A nonempty, closed and convex set $S \subseteq X$, with $S \neq X$, is said to be uniformly convex provided that

$$\delta_S(\epsilon) > 0, \quad \forall \epsilon \in \left\{ \begin{array}{ll} (0, \text{diam } S], & \text{if diam } S \text{ is attained on } S, \\ (0, \text{diam } S), & \text{otherwise.} \end{array} \right.$$
Remark 2.2. Since diam $S$ vanishes if $S$ is a singleton, Definition 2.1 (ii) does not exclude such kind of convex sets. Nevertheless, as singletons are not of 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 are a paradigm for the notion of uniform convexity for sets. Let us recall that, after [8], a Banach space is said to be uniformly convex if

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

Function $\delta_X$ is called modulus of convexity of the space $(X, \| \cdot \|)$. In fact it is possible to prove that

$$\delta_S(\varepsilon) = \delta_X(\varepsilon), \quad \forall \varepsilon \in (0, 2].$$

Such Banach spaces as $l^p, L^p$, and $W^p_n$, with $1 < p < \infty$, are all examples of uniformly convex spaces. In particular, every Hilbert space is uniformly convex. To the contrary, $c_0$, $L^1$ and $L^\infty$ are not so. Exactly known moduli of convexity for Banach spaces are, for instance,

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

provided that $p \geq 2$. For more details on uniformly convex Banach spaces and properties of their moduli the reader may refer to [8, 12, 17].

(ii) Uniformly convex sets can live also in Banach spaces that are not uniformly convex. Consider, for instance, given $r > 0$, the set $S = \{ x \in \mathbb{R}^n : \left( \sum_{i=1}^n x_i^2 \right)^{1/2} \leq r \}$ as a subset of the space $\mathbb{R}^n$, here equipped with the $\infty$-norm. The space $(\mathbb{R}^n, \| \cdot \|_\infty)$ is not uniformly convex but $S$ is uniformly convex. Moreover, if $n = 2$, by means of elementary geometric arguments, it is possible to determine the following expression for the modulus of convexity of $S$

$$\delta_S(\varepsilon) = \frac{1}{\sqrt{2}} \left( r - \sqrt{r^2 - \left( \frac{\varepsilon}{2} \right)^2} \right), \quad \forall \varepsilon \in (0, 2r].$$

Thus, it results in

$$\delta_S(\varepsilon) \geq \frac{\varepsilon^2}{8\sqrt{2r}}, \quad \forall \varepsilon \in (0, 2r].$$

Remark 2.4. (i) In [4] it has been established that every uniformly convex set is bounded. Directly from the definition, it follows that every uniformly convex set has nonempty interior. This fact entails that, while uniformly convex subsets of finite-dimensional spaces are compact, the same kind of sets can not be compact in infinite-dimensional Banach spaces.

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

$$\delta_S(\varepsilon) \leq a \varepsilon^2, \quad \forall \varepsilon \in (0, \text{diam } S)$$

(see [4]). Thus, a quadratic growth behaviour for a modulus of convexity is a maximal one.

The next proposition provides a complete characterization of uniform convexity for subsets of the Euclidean space $\mathbb{R}^n$ in terms of extremality of their boundary points.

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

Proof. Observe that by compactness of $S$, it is $\text{bd } S \neq \emptyset$. Actually, the Krein-Milman theorem ensures that $\text{ext } S \neq \emptyset$ also. Clearly, it is $\text{ext } S \subseteq \text{bd } S$. To begin with, assume that $S$ is uniformly convex. Take any $\bar{x} \in \text{bd } S$. If it were $\bar{x} \notin \text{ext } S$, then there would exist $x_1, x_2 \in S$, with $x_1 \neq x_2$, such that $\bar{x} = \frac{x_1 + x_2}{2}$. Consequently one would get $B(\bar{x}, \delta) \subseteq S$ for some $x_1, x_2 \in S$ only for $\delta = 0$, thereby contradicting the fact that, since $S$ is uniformly convex, it must be $\delta_S(\|x_1 - x_2\|) > 0$.

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

$$S_\varepsilon = \{ (x_1, x_2) \in S \times S : \|x_1 - x_2\| = \varepsilon \}$$
is still compact. Define the function \( \vartheta : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, +\infty) \) by setting
\[
\vartheta(x_1, x_2) = \text{dist} \left( \frac{x_1 + x_2}{2}, \mathbb{R}^n \setminus \text{int} S \right).
\]
Since such a function is continuous on \( \mathbb{R}^n \times \mathbb{R}^n \), it attains its global minimum over \( S^2 \) at some point \((\hat{x}_1, \hat{x}_2) \in S^2\). If it were \( \vartheta(\hat{x}_1, \hat{x}_2) = 0 \), then it would happen that
\[
\frac{\hat{x}_1 + \hat{x}_2}{2} \in \text{bd} \, S.
\]
The last inclusion contradicts the fact that \( \frac{\hat{x}_1 + \hat{x}_2}{2} \) is an extreme point of \( S \) (recall that \( x_1 \neq x_2 \)). Therefore, it must be \( \vartheta(\hat{x}_1, \hat{x}_2) > 0 \). As it is true that
\[
\delta_S(\epsilon) = \min_{(x_1, x_2) \in S^2} \vartheta(x_1, x_2) > 0,
\]
the requirement in Definition 2.1(ii) turns out to be satisfied. The arbitrariness of \( \epsilon \in (0, \text{diam} \, S] \) completes the proof.

**Remark 2.6.** (i) It is worth noting that one of the assertions in Proposition 2.5 remains true also in general Banach spaces. More precisely, if \( S \) is a uniformly convex subset of a Banach space, then each element of its boundary is an extreme point of \( S \). Of course, no compact convex subset of an infinite-dimensional Banach space can be uniformly convex.

(ii) From Definition 2.1 it follows that if \( S \) is uniformly convex and its modulus of convexity satisfies the quadratic growth condition
\[
\delta_S(\epsilon) \geq c\epsilon^2, \quad \forall \epsilon \in (0, \text{diam} \, S],
\]
for some \( c > 0 \), then for every \( \hat{c} \in (0, c) \) one has
\[
B \left( \frac{x_1 + x_2}{2}, \hat{c}\|x_1 - x_2\|^2 \right) \subseteq S, \quad \forall x_1, x_2 \in S.
\]

Another relevant class of convex subsets of a Banach space, which emerges in connection with the problem under study, can be obtained within the axiomatic approach to the notion of convexity (known as \( \Phi \)-convexity). This class was introduced in [19] and consists of all convex sets that are expressible as an intersection of balls with arbitrary but fixed radius and with center varying in an arbitrary subset of the underlying space, as formalized below.

**Definition 2.7.** Any subset \( S \subseteq X \) of the form
\[
S = \bigcap_{x \in M} B(x, r) \neq \emptyset,
\]
where \( r \) is a fixed positive real and \( M \subseteq X \) is an arbitrary subset, with \( M \neq X \), is called \( r \)-convex or strongly convex of radius \( r \).

Strongly convex sets and their properties have been studied by several authors (see [4, 5, 13, 20, 21, 32]). In the related literature the reader will find various characterizations of \( r \)-convexity. Some facts concerning \( r \)-convexity that are relevant to the present investigations are recalled below.

**Remark 2.8.** (i) It is readily seen that, if a Banach space \((X, \| \cdot \|)\) is uniformly convex with modulus \( \delta_X \), then any strongly convex set \( S \subseteq X \) with radius \( r \) is uniformly convex and its modulus of convexity satisfies the relation
\[
\delta_S(\epsilon) \geq r\delta_X \left( \frac{\epsilon}{r} \right), \quad \forall \epsilon \in (0, \text{diam} \, S].
\]

(ii) In [5] it has been established that the converse also is true, provided that \((X, \| \cdot \|)\) is a Hilbert space. In such setting, a subset \( S \) is uniformly convex if it is strongly convex for some radius \( r > 0 \). Note that, in the light of Proposition 2.5 a strongly convex subset of \( \mathbb{R}^n \) (equipped with its Euclidean structure) is characterized by the condition \( \text{ext} \, S = \text{bd} \, S \).

(iii) As a straightforward consequence of Definition 2.7 it follows that if \( S \) is \( r \)-convex, then \( \text{diam} \, S \leq 2r \).
ON A CLASS OF CONVEX SETS WITH CONVEX IMAGES AND ITS APPLICATION TO NONCONVEX OPTIMIZATION 5

3. Metric regularity on sets

In the next definition, the notion of metric regularity on sets introduced in [10], which was originally proposed for set-valued mappings, is adapted to the single-valued case. Let us start with considering this notion with reference to a pair \((X, d), (Y, d)\) of metric spaces.

**Definition 3.1.** Given two subsets \(U \subseteq X\) and \(V \subseteq Y\), a mapping \(f : X \rightarrow Y\) is said to be *metrically regular* on \(U\) for \(V\) if there is constant \(\kappa > 0\) such that

\[
\text{dist} (x, f^{-1}(y)) \leq \kappa \text{dist} (y, f(x)), \quad \forall x \in U : f(x) \in V, \forall y \in V.
\]

The value

\[
\text{reg}(f, U \times V) = \inf \{\kappa > 0 : \text{inequality (3.1) holds}\}
\]

is called *modulus of metric regularity* of \(f\) on \(U\) for \(V\).

As it happens for its local version, metric regularity on sets formalizes a property of mappings that has to do with a surjective behaviour of them. The next definition recalls such alternative description of regularity, in the case of interest.

**Definition 3.2.** Given two subsets \(U \subseteq X\) and \(V \subseteq Y\), a mapping \(f : X \rightarrow Y\) is said to be *linearly open* on \(U\) for \(V\) if there is constant \(\alpha > 0\) such that

\[
f(\text{int} B \{x, r\}) \supseteq [\text{int} B \{f(x), \alpha r\}] \cap V, \quad \forall r > 0, \forall x \in U : f(x) \in V.
\]

The value

\[
\text{cov}(f, U \times V) = \sup \{\alpha > 0 : \text{inclusion (3.2) holds}\}
\]

is called *modulus of linear openness* of \(f\) on \(U\) for \(V\).

The equivalence of these two notions is stated in the below proposition (see Proposition 1.5 in [10]).

**Proposition 3.3.** Let \(U\) and \(V\) be nonempty subsets of \(X\) and \(Y\), respectively, and let \(f : X \rightarrow Y\) be such that

\(g\text{ph} f \cap (U \times V) \neq \emptyset\).

Mapping \(f\) is metrically regular on \(U\) for \(V\) iff it is linearly open \(U\) for \(V\) and it holds

\[
\text{reg}(f, U \times V) = 1/\text{cov}(f, U \times V).
\]

Whenever it is \(V = Y\) in considering metric regularity/linear openness on sets, the following shortened notation will be used

\[
\text{reg}(f, U), \quad \text{cov}(f, U)
\]

in place of \(\text{reg}(f, U \times V)\) and \(\text{cov}(f, U \times Y)\), respectively.

The rest of the present section is devoted to formulating adequate conditions for metric regularity on sets. Following an approach introduced in [11] (see also [2, 3]), this task will be accomplished by making use of a coincidence point theorem. Recall that, given \(g : X \rightarrow Y\) and \(h : X \rightarrow Y\), an element \(x \in X\) is said to be a *coincidence point* of mappings \(g\) and \(h\) if

\[
g(x) = h(x).
\]

It is evident that such notion generalizes the notion of fixed point. In what follows, given a pair of nonempty set \(U \subseteq X\) and \(V \subseteq Y\), one will be concerned, in particular, with coincidence points in \(U\) having their images in \(V\), whose set is denoted by

\[
\text{Coin}(g, h, U \times V) = \{x \in U : g(x) = h(x) \in V\}.
\]

**Theorem 3.4.** Let \(g : X \rightarrow Y\) and \(h : X \rightarrow Y\) be mappings between metric spaces, and let \(U \subseteq X\) and \(V \subseteq Y\). Suppose that:

(i) \(U\) is metrically complete;

(ii) \(g\) is continuous and linearly open on \(U\) for \(V\);

(iii) \(h\) is Lipschitz on \(U\), with \(\text{lip}(h, U) < \text{cov}(g, U \times V)\);

(iv) \(g(U) \subseteq V\) and \(h(U) \subseteq V\).
Then $\text{Coin}(g, h, U \times V) \neq \emptyset$ and it holds
\[
\text{dist} \ (x, \text{Coin}(g, h, U \times V)) \leq \frac{d(g(x), h(x))}{\text{cov}(g, U \times V) - \text{lip}(h, U)}, \quad \forall x \in U.
\]

**Proof.** Fix an arbitrary $x_0 \in U$. If $g(x_0) = h(x_0)$, by hypothesis (iv) it is $x_0 \in \text{Coin}(g, h, U \times V)$, so that all assertions in the thesis are verified at once.

Otherwise, define the functional $\varphi : U \rightarrow [0, +\infty)$ as follows
\[
\varphi(x) = d(g(x), h(x)).
\]

Notice that, under the hypotheses assumed, $\varphi$ is continuous and bounded from below on $U$, whereas $U$ is metrically complete. So, taken an arbitrary $\eta \in (0, \text{cov}(g, U \times V) - \text{lip}(h, U))$, by virtue of the Ekeland’s variational principle, corresponding to
\[
\lambda = \frac{\varphi(x_0)}{\text{cov}(g, U \times V) - \text{lip}(h, U) - \eta},
\]
an element $x_\lambda \in U$ must exist such that
\[
d(x_\lambda, x_0) \leq \lambda
\]
and
\[
\varphi(x_\lambda) < \varphi(x) + [\text{cov}(g, U \times V) - \text{lip}(h, U) - \eta]d(x, x_\lambda), \quad \forall x \in U \setminus \{x_\lambda\}.
\]

It is possible to prove that $\varphi(x_\lambda) = 0$. Indeed, assume to the contrary that $\varphi(x_\lambda) > 0$. Since $h(x_\lambda)$ belongs to $V$, then, taken any $\zeta$ with
\[
0 < \zeta < \frac{\text{cov}(g, U \times V)}{\text{cov}(g, U \times V) - \eta} - 1,
\]
it results in
\[
h(x_\lambda) \in \text{int} B(g(x_\lambda), \varphi(x_\lambda)(1 + \zeta)).
\]

Consequently, by the linear openness of $g$ at $U$ for $V$, there exists $u \in \text{int} B\left(x_\lambda, \frac{\varphi(x_\lambda)(1 + \zeta)}{\text{cov}(g, U \times V)}\right) \setminus \{x_\lambda\}$ such that
\[
g(u) = h(x_\lambda).
\]

Thus, by the hypothesis (iii), the following estimate holds
\[
\varphi(u) = d(g(u), h(u)) \leq d(g(u), h(x_\lambda)) + d(h(x_\lambda), h(u)) \leq \text{lip}(h, U)d(x_\lambda, u).
\]

By taking into account the last estimate, from inequalities (3.3a) and (3.3b) one obtains
\[
\varphi(x_\lambda) < \varphi(u) + [\text{cov}(g, U \times V) - \text{lip}(h, U) - \eta]d(u, x_\lambda) \leq [\text{cov}(g, U \times V) - \eta]d(u, x_\lambda)
\]
\[
< \frac{\text{cov}(g, U \times V) - \eta}{\text{cov}(g, U \times V)} \varphi(x_\lambda)(1 + \zeta) < \varphi(x_\lambda),
\]

which is clearly inconsistent. This shows that $\varphi(x_\lambda) = 0$ and hence $x_\lambda \in \text{Coin}(g, h, U \times V)$. Moreover, by recalling inequality (3.4) and the value of $\lambda$, as chosen in (3.3), one finds
\[
\text{dist} \ (x_0, \text{Coin}(g, h, U \times V)) \leq d(x_0, x_\lambda) \leq \frac{\varphi(x_0)}{\text{cov}(g, U \times V) - \text{lip}(h, U) - \eta}.
\]

The arbitrariness of $\eta \in (0, \text{cov}(g, U \times V) - \text{lip}(h, U))$ allows one to obtain the inequality to be proved, in the case $x = x_0$. Being $x_0$ arbitrary, the proof is completely proved. \[\square\]

Let us assume now $(Y, d)$ to be a metric linear space, whose metric is invariant under translation. In such a setting, Theorem 3.4 pave the way to extending to the metric regularity of sets the well-known Milyutin’s theorem on the stability of linear openness in the presence of Lipschitz perturbations (see [11], [12], [13]).
Proposition 3.5. Let $g: X \to Y$ and $l: X \to Y$ be mappings between metric spaces, and let $U \subseteq X$ and $V \subseteq Y$ be nonempty sets. Suppose that:

(i) $U$ is metrically complete;
(ii) $g$ is continuous and linearly open on $U$ for $V$;
(iii) $l$ is Lipschitz on $U$, with $\text{lip}(l, U) < \text{cov}(g, U \times V)$;
(iv) $g(U) \subseteq V$ and $V - l(U) \subseteq V$.

Then it holds

$$\text{cov}(g + l, U \times V) \geq \text{cov}(g, U \times V) - \text{lip}(l, U)$$

and hence also $g + l$ is linearly open on $U$ for $V$.

Proof. To prove the thesis, one has to show that for every $\alpha < \text{cov}(g, U \times V) - \text{lip}(l, U)$, the following inclusion holds true

$$(g + l)(\text{int} B(x, r)) \supseteq \text{int} B(g(x) + l(x), \alpha r) \cap V, \quad \forall r > 0, \forall x \in U : g(x) + l(x) \in V.$$ 

So, fix $x_0 \in U$ and $r > 0$, and take $\alpha \in (0, \text{cov}(g, U \times V) - \text{lip}(l, U))$. Let us consider now an arbitrary $y \in [\text{int} B(g(x_0) + l(x_0), \alpha r)] \cap V$. Define the mapping $h: X \to Y$ as $h = y - l$. By hypothesis (iv), one has

$$h(U) \subseteq y - l(U) \subseteq V - l(U) \subseteq V.$$ 

Besides, by the invariance under translation of the metric of $Y$, it is $\text{lip}(h, U) = \text{lip}(l, U)$. It is then possible to apply Theorem 3.4, according to which, corresponding to $\eta$ such that $d(y, g(x_0) + l(x_0))(1 + \eta) < \alpha r$, there exists $\bar{x} \in U$ with the properties

$$g(\bar{x}) = y - l(\bar{x})$$

and

$$d(\bar{x}, x_0) \leq \frac{d(g(x_0), y - l(x_0))}{\text{cov}(g, U \times V) - \text{lip}(h, U)} (1 + \eta).$$

Thus it follows that

$$y = (g + l)(\bar{x})$$

and

$$d(\bar{x}, x_0) \leq \frac{d(y, g(x_0) + l(x_0))}{\text{cov}(g, U \times V) - \text{lip}(h, U)} (1 + \eta) \leq \frac{\alpha r}{\text{cov}(g, U \times V) - \text{lip}(h, U)} < r.$$ 

The last inequality means that $\bar{x} \in \text{int} B(x_0, r)$, so inclusion (3.7) is proved to hold true. This completes the proof. 

Henceforth $(\mathbb{X},\|\cdot\|)$ and $(\mathbb{Y},\|\cdot\|)$ will denote real Banach spaces. The persistence of metric regularity on sets under perturbations can be exploited to establish sufficient conditions for detecting such property by means of simple approximations of the original mapping. Here, as an approximation of a mapping, a slight modification of the notion of strict estimator (see [11]) is considered, according to the below definition.

Definition 3.6. Let $f: \mathbb{X} \to \mathbb{Y}$ be a mapping, let $U \subseteq \mathbb{X}$ be a nonempty set, and let $\epsilon$ be a positive real. A mapping $g: \mathbb{X} \to \mathbb{Y}$ is said to be a strict $\epsilon$-variation of $f$ on $U$ if

$$\text{lip}(f - g, U) \leq \epsilon.$$ 

Below an example is proposed which will be useful in subsequent developments.

Example 3.7. Let $f: \mathbb{R} \to \mathbb{R}$, let $U \subseteq \mathbb{R}$ be a nonempty, closed and convex set, and let $\Omega \subseteq \mathbb{R}$ be an open subset such that $\Omega \supseteq U$. Assume that $f \in C^{1,1}(\Omega)$. Then, for any $\bar{x} \in U$, mapping $D_f(\bar{x})$ is a strict $\epsilon$-variation of $f$ on $U$, for every $\epsilon \geq \text{lip}(D_f, U)$. Indeed, by applying the mean value theorem to the mapping $f - D_f(\bar{x})[\cdot - \bar{x}]$, one finds

$$\|f(x_1) - f(x_2) - D_f(\bar{x})[x_1 - x_2]\| \leq \sup_{x \in [x_1, x_2]} \|D_f(x) - D_f(\bar{x})\| \|x_1 - x_2\|$$

$$\leq \text{lip}(D_f, U)\|x_1 - x_2\|, \quad \forall x_1, x_2 \in U.$$ 

Thus $\text{lip}(f - D_f(\bar{x})[\cdot - \bar{x}], U) \leq \text{lip}(D_f, U) \leq \epsilon$.

In terms of strict $\epsilon$-variations, the Milyutin’s theorem for regularity on sets can be conveniently reformulated as follows.
Proposition 3.8. Let \( f : X \rightarrow Y \) and \( g : X \rightarrow Y \) be mappings between Banach spaces, and let \( U \subseteq X \) be a nonempty closed set. Suppose that:

(i) \( g \) is continuous and linearly open on \( U \);
(ii) \( g \) is a strict \( \epsilon \)-variation of \( f \) on \( U \), with \( \epsilon < \text{cov}(g,U) \).

Then also \( f \) is linearly open on \( U \) and it holds

\[
\text{cov}(f,U) \geq \text{cov}(g,U) - \epsilon.
\]

Proof. In view of applying Proposition 3.5, define \( l = f - g \). Observe that \( U \) is metrically complete, as a closed subset of a Banach space. Since in this case \( V = Y \), it is automatically true that

\[
g(U) \subseteq Y \quad \text{and} \quad Y - l(U) \subseteq Y.
\]

Moreover, by hypotheses (ii), it is \( \text{lip}(l,U) \leq \epsilon < \text{cov}(g,U) \). Then the thesis follows at once by Proposition 3.5 with

\[
\text{cov}(f,U) = \text{cov}(g + f - g, U) \geq \text{cov}(g,U) - \text{lip}(f-g, U) \geq \text{cov}(g,U) - \epsilon.
\]

\[\square\]

For the purposes of the present analysis, conditions for the metric regularity on sets of smooth mappings are needed. In the case of strictly differentiable mappings between Banach spaces, the main criterion characterizing the regularity around a point is the Lyusternik-Graves theorem (see [10, 11, 18]). In order to establish a similar result for metric regularity on sets, it is useful to recall the notion of dual Banach constant of a linear operator.

By such term the quantity

\[
b^*(h) = \inf_{v^* \in E} \|h^*[v^*]\|_*
\]

is meant, where \( h^* \in \mathcal{L}(Y^*, X^*) \) stands for the adjoint operator to \( h \in \mathcal{L}(X, Y) \). The dual Banach constant appears in the quantitative formulation of the Banach-Schauder theorem, according to which \( h \in \mathcal{L}(X, Y) \) is linearly open on \( X \) (for \( Y \)) iff \( b^*(h) > 0 \). Actually, in such an event one has \( \text{cov}(h,U) = b^*(h) \), for every \( U \subseteq X \).

Given a mapping \( f : X \rightarrow Y \), with \( f \in C^{1,1}(\Omega) \) and a nonempty set \( U \subseteq \Omega \), define

\[
b^*(Df, U) = \sup_{x \in U} b^*(Df(x)).
\]

Corollary 3.9. Let \( f : X \rightarrow Y \), let \( U \subseteq X \) be a nonempty, closed and convex set, and let \( \Omega \subseteq X \) be an open set such that \( \Omega \supseteq U \). If \( f \in C^{1,1}(\Omega) \) and \( b^*(Df, U) > \text{lip}(Df, U) \), then \( f \) is linearly open/metrically regular on \( U \), and it results in

\[
\text{cov}(f,U) \geq b^*(Df, U) - \text{lip}(Df, U).
\]

Proof. Let \( \eta \) be an arbitrary positive real in \((0, b^*(Df, U) - \text{lip}(Df, U))\). Correspondingly, by the definition of \( b^*(Df, U) \) an element \( x_\eta \) can be found in \( U \) such that \( b^*(Df(x_\eta)) > b^*(Df,U) - \eta \). Since, under the current assumptions, \( f \) admits the mapping \( Df(x_\eta) \) as a strict \( \text{lip}(Df, U) \)-variation on \( U \) (recall Example 3.7), it suffices to apply Proposition 3.5 in order to get

\[
\text{cov}(f,U) \geq \text{cov}(Df(x_\eta), U) - \text{lip}(Df, U) = b^*(Df(x_\eta)) - \text{lip}(Df, U) > b^*(Df, U) - \text{lip}(Df, U) - \eta,
\]

wherefrom, by arbitrariness of \( \eta \), the thesis follows.

\[\square\]

Because the sufficient condition coming up from Corollary 3.9 is interesting not only as a mere tool of analysis, an example of a concrete circumstance, in which it can be employed, is proposed below.

Example 3.10. Let \( X = Y = \mathbb{R} \) be equipped with its usual Euclidean structure. Consider the function \( f : \mathbb{R} \rightarrow \mathbb{R} \) defined by

\[
f(t) = \arctan t.
\]

Clearly, it is \( f \in C^\infty(\mathbb{R}) \) and, as it is well known, it results in

\[
Df(t) = \frac{1}{1 + t^2}, \quad D^2f(t) = -\frac{2t}{(1 + t^2)^2},
\]

where \( D^k f(t) \) denotes the \( k \)-order derivative of \( f \) at \( t \). By considering the zeros and the sign of

\[
D^3f(t) = -\frac{2(1 - 3t^2)}{(1 + t^2)^3},
\]
one sees that function \( t \mapsto |D^2f(t)| \) attains its (global) maximum value for \( t \in \{-\sqrt{3}, \sqrt{3}\} \). Therefore, in the case under consideration, one finds

\[
b^*(Df, \mathbb{R}) = \sup_{t \in \mathbb{R}} b^*(Df(t)) = \sup_{t \in \mathbb{R}} |Df(t)| = |Df(0)| = 1.
\]

On the other hand, the mean value theorem allows one to get the following estimate

\[
\operatorname{lip}(Df, \mathbb{R}) \leq \sup_{t \in \mathbb{R}} |D^2f(t)| = \left| D^2f \left( \pm \frac{1}{\sqrt{3}} \right) \right| = \frac{3\sqrt{3}}{8}.
\]

According to Corollary 3.9, since for each subset \( U \) containing 0 it is true that

\[
b^*(Df, U) = \sup_{t \in U} b^*(Df(t)) = \sup_{t \in U} |Df(t)| = |Df(0)| = 1 > \frac{3\sqrt{3}}{8} \geq \operatorname{lip}(Df, U),
\]

\( f \) turns out to be linearly open on each \( U \) and the following global estimate is then established

\[
\operatorname{cov}(f, U) \geq 1 - \frac{3\sqrt{3}}{8}.
\]

Of course, the sufficient condition for metric regularity on sets provided by Corollary 3.9 is specific for \( C^{1,1} \) mappings. A more general and deeper treatment of this topic can be found in [16].

As a further tool of analysis, the below estimate will be crucially employed in proving the main result.

**Lemma 3.11.** Let \( f : X \to Y \) be a mapping between Banach spaces, let \( U \subseteq X \), let \( \Omega \subseteq X \) be an open set such that \( \Omega \supseteq U \), and let \( x_1, x_2 \in U \), with \( [x_1, x_2] \subseteq U \). If \( f \in C^{1,1}(\Omega) \), then it holds

\[
\left\| \frac{f(x_1) + f(x_2)}{2} - f \left( \frac{x_1 + x_2}{2} \right) \right\| \leq \frac{\operatorname{lip}(Df, U)}{8} \|x_1 - x_2\|^2.
\]

Its proof can be found, for instance, in [30] (see Lemma 2.7).

### 4. Convexity of Images through Smooth Regular Mappings

The main result of the paper is the following sufficient condition for the convexity of the images of uniformly convex sets through \( C^{1,1} \) mappings between general Banach spaces.

**Theorem 4.1.** Let \( f : X \to Y \) be a mapping between Banach spaces, let \( S \subseteq X \) be a closed set and let \( \Omega \subseteq X \) be an open set such that \( \Omega \supseteq S \). Suppose that:

(i) \( S \) is uniformly convex with modulus fulfilling the quadratic growth condition (2.1), for a proper \( c > 0 \);

(ii) \( f \in C^{1,1}(\Omega) \) and \( b^*(Df, S) > \operatorname{lip}(Df, S) \);

(iii) it holds

\[
\frac{\operatorname{lip}(Df, S)}{8|b^*(Df, S) - \operatorname{lip}(Df, S)|} < c; \tag{4.1}
\]

(iv) \( f(S) \) is closed.

Then \( f(S) \) is convex.

**Proof.** Since \( f(S) \) is a closed set by hypothesis, it suffices to show that, for any arbitrary pair \( y_1, y_2 \in f(S) \), it happens that

\[
y = \frac{y_1 + y_2}{2} \in f(S).
\]

Take \( x_1, x_2 \in S \) in such a way that

\[
y_1 = f(x_1) \quad \text{and} \quad y_2 = f(x_2),
\]

and define

\[
x = \frac{x_1 + x_2}{2}.
\]
Since it is \( x \in S \), if it happens that \( f(x) = \tilde{y} \), one can immediately conclude that \( \tilde{y} \in f(S) \). Otherwise, notice that, by virtue of Corollary 3.9 under the above assumptions \( f \) turns out to be metrically regular on \( S \), with the following modulus of regularity

\[
\text{reg}(f, S) \leq \frac{1}{b^*(Df, S) - \text{lip}(Df, S)}
\]

(recall Proposition 3.3). This fact implies in particular that

\[
\text{dist}(\bar{x}, f^{-1}(\bar{y})) \leq \frac{\|\bar{y} - f(\bar{x})\|}{b^*(Df, S) - \text{lip}(Df, S)}
\]

By exploiting the estimate provided by Lemma 3.11 from the last inequality, taking any \( \eta \) with

\[
0 < \eta < \frac{8c[b^*(Df, S) - \text{lip}(Df, S)]}{\text{lip}(Df, S)} - 1,
\]

one gets the existence of \( \hat{x} \in f^{-1}(\bar{y}) \) such that

\[
d(\hat{x}, \bar{x}) \leq \frac{\|\tilde{y} - f(\bar{x})\|}{b^*(Df, S) - \text{lip}(Df, S)} \left(1 + \eta\right) \leq \frac{\text{lip}(Df, S)\|x_1 - x_2\|^2}{8b^*(Df, S) - \text{lip}(Df, S)} \left(1 + \eta\right).
\]

On account of hypothesis (iv), from the last inequality chain one obtains

\[
d(\hat{x}, \bar{x}) < c\|x_1 - x_2\|^2.
\]

Since by hypothesis \( S \) is uniformly convex and its modulus of convexity satisfies condition (2.3), then, according to Remark 2.6 (ii), for any \( \tilde{c} \in (0, c) \) it results in \( B(\bar{x}, \tilde{c}\|x_1 - x_2\|^2) \subseteq S \). Therefore it follows that \( \hat{x} \in S \) and hence \( \tilde{y} \in f(S) \). This completes the proof.

\[\square\]

**Remark 4.2.** (i) Since as a consequence of hypothesis (ii) of Theorem 4.1 mapping \( f \) is continuous on \( S \), then in finite-dimensional spaces hypothesis (iv) is automatically satisfied.

(ii) A feature of Theorem 4.1 to be noticed is that it makes no direct assumption on the Banach space \( (X, \|\cdot\|) \). The geometric requirement behind the result is made instead on the set \( S \). In the light of Example 2.3 (ii), such a feature considerably enlarges the range of applicability of the original Polyak’s convexity principle. In fact, an “at point formulation” is replaced with a non-local “on set formulation”.

In a certain subclass of uniformly convex Banach spaces it is possible to establish an analogous result for the class of \( r \)-convex sets.

**Theorem 4.3.** Let \( f : X \rightarrow Y \) be a mapping between Banach spaces, let \( S \subseteq X \) be a closed set and let \( \Omega \subseteq X \) be an open set such that \( \Omega \supseteq S \). Suppose that:

(i) \( (X, \|\cdot\|) \) is uniformly convex with a modulus of convexity \( \delta_X \) satisfying the quadratic growth condition

\[
\delta_X(\epsilon) \geq \kappa \epsilon^2, \quad \forall \epsilon \in (0, 2];
\]

(ii) \( f \in C^{1,1}(\Omega) \) and \( b^*(Df, S) > \text{lip}(Df, S) \);

(iii) \( S \) is strongly convex with a radius \( r \) such that

\[
r < \frac{8\kappa[b^*(Df, S) - \text{lip}(Df, S)]}{\text{lip}(Df, S)} \tag{4.2}
\]

(iv) \( f(S) \) is closed.

Then \( f(S) \) is convex.

**Proof.** In force of hypothesis (i), according to Remark 2.8 (i), the set \( S \) turns out to be uniformly convex with modulus

\[
\delta_S(\epsilon) \geq r \kappa \left(\frac{\epsilon}{r}\right)^2 = \frac{\kappa \epsilon^2}{r^2}.
\]

It remains to apply Theorem 4.1 with \( c = \kappa/r \), where condition (4.1) becomes (4.2). \[\square\]
ON A CLASS OF CONVEX SETS WITH CONVEX IMAGES AND ITS APPLICATION TO NONCONVEX OPTIMIZATION

**Remark 4.4.** (i) Hypothesis (i) of Theorem 4.3 is known to be satisfied for all such Banach spaces as $L^p$, $L^q$, and $W^p_m$, with $1 < p < 2$ (see [12 17]). Of course, it is also satisfied in a Hilbert space (remember Example 2.3 (i)).

(ii) It is useful to note that in a Hilbert space setting condition (4.2) takes the simpler form

$$r < \frac{b^*(Df, S)}{\text{lip}(Df, S)} - 1,$$

in as much as in such case one can take $\kappa = 1/8$. Notice that Theorem 2.1 in [24] can not be directly derived from the last result. This is due to the fact that the modulus of linear openness on $S$ is given here by $b^*(Df, S) - \text{lip}(Df, S)$. Nonetheless if, for a given $x_0 \in \Omega$, $Df(x_0)$ is onto, then, by taking a ball around $x_0$ with radius small enough, the Lyusternik-Graves theorem allows one to replace $b^*(Df, S) - \text{lip}(Df, S)$ with $b^*(Df(x_0))$

Consequently, condition (4.2) becomes

$$r \text{lip}(Df, B(x_0, r)) < b^*(Df(x_0))$$

and the positivity of the difference $b^*(Df, S) - \text{lip}(Df, S)$ is no longer required. Thus, in order to recover Theorem 2.1 in [24], it suffices to take $r$ sufficiently small.

(iii) A result comparable with Theorem 4.3 has been recently presented in [28] (see, in particular, Proposition IV.1). It states the convexity of images of $r$-convex sets through $C^{1,1}$-diffeomorphisms between open subsets of $\mathbb{R}^n$. The key condition upon which this result has been achieved relates to the geometric behaviour of normals to boundary points of the reference set.

To illustrate the above theory, let us consider its employment in dealing with (homogeneous) quadratic transformations between finite-dimensional Euclidean spaces $q : \mathbb{R}^n \rightarrow \mathbb{R}^m$, given by

$$q(x) = \begin{pmatrix} q_1(x) \\ \vdots \\ q_m(x) \end{pmatrix}, \quad x \in \mathbb{R}^n,$$

where each $q_i(x) = x^\top Q_i x$, for $i = 1, \ldots, m$, is a quadratic form on $\mathbb{R}^n$, represented by a symmetric $n \times n$ matrix $Q_i$ with real entries (vectors of $\mathbb{R}^n$ are treated here as columns). If $\mathbb{R}^n$ is equipped with its usual Euclidean space structure what noted in Remark 4.4 (ii) applies. Of course, it is $q \in C^2(\mathbb{R}^n)$ and, through trivial calculations, it is readily seen that

$$Dq(x) = 2 \begin{pmatrix} x^\top Q_1 \\ \vdots \\ x^\top Q_m \end{pmatrix}, \quad x \in \mathbb{R}^n.$$

If denoting by $(\mathcal{M}_{m \times n}(\mathbb{R}), \| \cdot \|_\mathcal{M})$ the space of $m \times n$ matrices with real entries, endowed with the 1-norm, the mapping $Dq : \mathbb{R}^n \rightarrow \mathcal{M}_{m \times n}(\mathbb{R})$ turns out to be linear. As such, it is Lipschitz continuous on $\mathbb{R}^n$ and its Lipschitz modulus amounts to

$$\text{lip}(Dq, \mathbb{R}^n) = \|Dq\|_\mathcal{L} = 2 \max_{u \in \mathbb{S}} \left\| \begin{pmatrix} u^\top Q_1 \\ \vdots \\ u^\top Q_m \end{pmatrix} \right\|_\mathcal{M}.$$

Now, fix $x \in \mathbb{R}^n$. Since each matrix $Q_i$ is symmetric and $(\mathbb{R}^n)^* \cong \mathbb{R}^n$, one has

$$b^*(Dq(x)) = 2 \min_{v \in \mathbb{S}} \| (Q_1 x, \ldots, Q_m x) v \|.$$

Given a set $S \subseteq \mathbb{R}^n$, the reader should notice that, unlike $\text{lip}(Dq, S)$, in the case under consideration $b^*(Dq, S)$ depends essentially on $S$. According to Theorem 4.3 whenever a strongly convex set $S$ is such that

$$\max_{x \in S} \min_{v \in \mathbb{S}} \| (Q_1 x, \ldots, Q_m x) v \| > \max_{u \in \mathbb{S}} \left\| \begin{pmatrix} u^\top Q_1 \\ \vdots \\ u^\top Q_m \end{pmatrix} \right\|_\mathcal{M},$$

11
the quadratic mapping \( q \) preserves the convexity of \( S \), provided that the radius \( r \) of \( S \) fulfills the condition

\[
(4.5) \quad r < \max_{x \in S} \min_{v \in S} \| (Q_1 x, \ldots, Q_m x) v \| \cdot \left( \max_{u \in S} \left\| \begin{pmatrix} u^T Q_1 \\ \vdots \\ u^T Q_m \end{pmatrix} \right\|_\mathcal{M} \right)^{-1} - 1.
\]

**Example 4.5.** Letting \( n = m = 2 \), consider the quadratic transformation \( q : \mathbb{R}^2 \to \mathbb{R}^2 \) defined by

\[
q(x) = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1^2 - x_2^2 \end{pmatrix}, \quad x \in \mathbb{R}^2.
\]

After trivial calculations, one finds

\[
Dq(0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},
\]

so that \( Dq(0) \) fails to be onto. On the base of the Lyusternik-Graves theorem this is necessary and sufficient in order for \( q \) to be not metrically regular around \( 0 \). As a consequence, the Polyak's convexity principle cannot be invoked, if taking \( 0 \) as a reference point. Nevertheless, by direct inspection one can observe that \( q \) transforms balls centered at \( 0 \), with arbitrarily large radius, into convex sets. Indeed, if passing to the polar coordinate system \((t, \rho) \in [0, 2\pi) \times [0, \infty)\), the above quadratic mapping takes the following form

\[
q(t, \rho) = \begin{pmatrix} \rho^2 (t - \sin^2 t) \\ \rho^2 (\cos^2 t - \sin^2 t) \end{pmatrix}.
\]

Thus, since setting \( y_2(t) = \cos^2 t - \sin^2 t \) it is \( y_2([0, 2\pi]) = [-1, 1] \), it results in

\[
q(B(0, r)) = \{ y = (y_1, y_2) \in \mathbb{R}^2 : y_1 \in [0, r^2], |y_2| \leq y_1 \}.
\]

So, for every \( r > 0 \) the image of \( B(0, r) \) through \( q \) is a convex set. As one expects, the Polyak's convexity principle is far removed from providing a necessary condition for the convexity of images.

Let us use this example to test the above theory. In the current case one has

\[
Q_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

Therefore, one obtains

\[
\max_{u \in S} \left\| \begin{pmatrix} u^T Q_1 \\ u^T Q_2 \end{pmatrix} \right\|_\mathcal{M} = \max_{t \in [0, 2\pi]} \left\| \begin{pmatrix} \cos t & \sin t \\ \cos t & -\sin t \end{pmatrix} \right\|_\mathcal{M} = \max_{t \in [0, 2\pi]} 2[|\cos t| + |\sin t|] = 2\sqrt{2}.
\]

On the other hand, for any \( x \in \mathbb{R}^2 \) it results in

\[
\min_{v \in S} \|(Q_1 x, Q_2 x) v\| = \min_{t \in [0, 2\pi]} \left\| \begin{pmatrix} x_1 (\cos t + \sin t) \\ x_2 (\cos t - \sin t) \end{pmatrix} \right\| = \min_{t \in [0, 2\pi]} \left[ x_1^2 (\cos t + \sin t)^2 + x_2^2 (\cos t - \sin t)^2 \right]^{1/2} \leq \|x\|.
\]

It follows that

\[
\max_{x \in B(0, r)} \min_{v \in S} \|(Q_1 x, Q_2 x) v\| \leq r
\]

and hence, for any \( r \) with \( 0 < r < 2\sqrt{2} \), condition (4.3) happens here to be violated. However, the same quadratic transformation allows one also to show a case in which, by changing reference set, conditions (4.3) and (4.5) may be both fulfilled. This occurs, for instance, if a strongly convex subset of \( \mathbb{R}^2 \) meets the bisectrix of the nonnegative orthant \( \mathbb{R}_+^2 \) sufficiently away from the origin. In this way, it is possible to construct \( r \)-convex sets preserving their convexity through the given mapping, with arbitrarily large radius. Indeed, if \( x_\xi = (\xi, \xi)^T \in S \), with \( \xi > 0 \), one finds

\[
\min_{v \in S} \|(Q_1 x_\xi, Q_2 x_\xi) v\| = \xi \sqrt{2}.
\]

Therefore, (4.4) and (4.5) are both fulfilled if \( \xi > 2 \) and \( r < \frac{\xi}{2} - 1 \).
ON A CLASS OF CONVEX SETS WITH CONVEX IMAGES AND ITS APPLICATION TO NONCONVEX OPTIMIZATION

5. Convexity in nonconvex optimization

In the previous sections a geometric problem dealing with abstract sets and mappings has been considered and discussed. The aim of this section is to expose some applications of the main results achieved to nonconvex optimization problems. By the latter term, any constrained extremum problem is meant in which either the feasible region or the objective functional (possibly both) is not necessarily convex. In handling this kind of problems, the lack of such a fundamental property as convexity does not allow one to take profit from basic principles of convex analysis. These are known to guarantee to optimization problems high desirable features such as globality of solutions, zero-order Lagrangian type optimality conditions, the fact that necessary conditions become often full characterizations of optimality. Yet, as the Polyak’s convexity reveals, a certain amount of convexity may lie hidden even in nonconvex optimization problems. This convexity can emerge when considering images of uniformly convex or strongly convex sets through $C^{1,1}$ mappings associated with the problems. This useful phenomenon is here exploited in two different contexts: solution existence and optimality characterization.

The convex behaviour of some classes of nonlinear optimization problems has been observed and investigated already in $[23, 24, 25, 26, 27, 30, 31]$. In all such references, based on the Polyak’s convexity principle, such a behaviour has been established to hold locally (whence the term “local programming”). The novelty of the results here presented consists in establishing a convex behaviour in a nonlocal form.

5.1. Global solution existence in constrained optimization. A rich variety of infinite-dimensional constrained optimization problems can be formalized as

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

where $\varphi : X \to \mathbb{R}$ describes the object (or cost) functional, while the constraints are expressed in a mixed form by a subset $S \subseteq X$ (geometric constraint) and by an inclusion $g(x) \in C$ (functional constraint), where $g : X \to Y$ is a given mapping and $C \subseteq Y$ is a nonempty closed set. In other terms, the feasible region of $(\mathcal{P})$ is

$$\mathcal{R} = S \cap g^{-1}(C).$$

Throughout the present subsection $\mathcal{R}$ will be assumed to be nonempty. If assuming $\varphi$ to be (at least) lower semicontinuous on the Banach space $(X, \| \cdot \|)$, a way to guarantee existence of global solution to $(\mathcal{P})$ is to introduce a compactness assumption on $\mathcal{R}$ or on its intersections with the sublevel sets of $\varphi$. Since such an assumption is very demanding in infinite-dimensional spaces, $(X, \| \cdot \|)$ is typically required to be reflexive. In such a setting, weak compactness is characterized in terms of boundedness and weak closedness, so that, as for existence questions, it becomes crucial to pass from norm closedness/continuity to weak closedness/continuity. In the light of the well-known Mazur’s theorem, a key assumption allowing one to do so is convexity of the feasible region and (quasi)convexity of the objective functional. Classical results for solution existence in constrained optimization problems essentially stem from a similar scheme. Notice that, with reference to problem $(\mathcal{P})$, the convexity of $\mathcal{R}$ may be lost even though $S$ is convex, owing to the presence of the functional constraint, which may drastically change the geometry of $\mathcal{R}$. Below an existence result is presented, which gives up the convexity of $\mathcal{R}$ and the quasiconvexity of $\varphi$. To see how this can be obtained, fix any $x_0 \in \mathcal{R}$ and define the mapping $I_{\mathcal{P}, x_0} : X \to \mathbb{R} \times Y$ as

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

and the set $Q = (-\infty, 0) \times C \subseteq \mathbb{R} \times Y$. Sets $S$, $Q$ and mapping $I_{\mathcal{P}, x_0}$ contain the essential informations needed to analyze problem $(\mathcal{P})$, according to a well-known approach to constrained optimization (see, for instance, $[9, 14, 15]$). It is clear that $x_0 \in \mathcal{R}$ is a global solution to $(\mathcal{P})$ iff

$$I_{\mathcal{P}, x_0}(S) \cap Q = \emptyset.$$

Suppose now that $\varphi, g \in C^{1,1}(X)$. Then it is possible to associate with $(\mathcal{P})$ the following two quantities

$$b_\mathcal{P}^* = b^*(D I_{\mathcal{P}, x_0}, S) \quad \text{and} \quad \text{lip}_\mathcal{P} = \text{lip}(D I_{\mathcal{P}, x_0}, S).$$

Notice that both the above numbers depend on the problem data $\varphi$, $g$ and $S$, whereas they are independent of $x_0$. 
Theorem 5.1. With reference to problem (P), suppose that:
(i) S is uniformly convex with modulus of convexity fulfilling condition \( \mu \), for some \( c > 0 \);
(ii) the Banach space \((Y, \| \cdot \|)\) is reflexive;
(iii) \( C \) is closed and convex;
(iv) for some \( x_0 \in \mathcal{R} \), set \( \mathcal{I}_{P,x_0}(S) \) is closed and bounded;
(v) \( \varphi, g \in C^{1,1}(\Xi) \) and it holds
\[
\text{b}_P^* > \text{lip}_P \quad \text{and} \quad \frac{\text{lip}_P}{8(\text{b}_P^* - \text{lip}_P)} < c.
\]
Then (P) admits a global solution.

Proof. Take an arbitrary \( x_0 \in \mathcal{R} \) and suppose that it is not a global solution to (P). As already remarked, this fact is equivalent to
\[
\mathcal{I}_{P,x_0}(S) \cap Q = \emptyset.
\]
Under the above hypotheses, it is possible to apply Theorem 4.1 to mapping \( \mathcal{I}_{P,x_0} \), in such a way to obtain that \( \mathcal{I}_{P,x_0}(S) \) is a convex subset of \( \mathbb{R} \times Y \). Now define
\[
\tau = \inf \{ t : (t, y) \in \mathcal{I}_{P,x_0}(S) \cap Q \}.
\]
On account of (5.1), the above infimum is taken over a nonempty set and, as it can be readily seen, actually one has
\[
\tau = \inf \{ t : (t, y) \in \mathcal{I}_{P,x_0}(S) \cap ((-\infty, 0] \times C) \}.
\]
Observe that the set \( \mathcal{I}_{P,x_0}(S) \cap ((-\infty, 0] \times C) \) is convex and closed, by virtue of the hypotheses (iii) and (iv) (actually, each set \( \mathcal{I}_{P,x}(S) \), with \( x \in \Xi \), is closed, being a translation of \( \mathcal{I}_{P,x_0}(S) \)). Thus, it turns out to be weakly closed. Since it is also bounded, the reflexivity of \((Y, \| \cdot \|)\) implies that \( \mathcal{I}_{P,x_0}(S) \cap ((-\infty, 0] \times C) \) is weakly compact. On the other hand, the projection operator \( \Pi_\mathcal{R} : \mathbb{R} \times Y \rightarrow \mathbb{R} \) is continuous and convex and hence it is also weakly lower semicontinuous. This allows one to assert that the infimum appearing in (5.2) is actually attained at some element of \( \mathcal{I}_{P,x_0}(S) \cap ((-\infty, 0] \times C) \), say \((\bar{t}, \bar{y})\). The existence of such an element entails the existence of \( \bar{x} \in S \) such that
\[
\bar{t} = \tau = \varphi(\bar{x}) - \varphi(x_0) < 0, \quad \bar{y} = g(\bar{x}) \in C.
\]
Let us show that \( \bar{x} \) is a global solution to (P). As it is true that \( \bar{x} \in S \cap g^{-1}(C) \), one has \( \bar{x} \in \mathcal{R} \). Now, suppose that there is \( \hat{x} \in \mathcal{R} \) such that
\[
\varphi(\hat{x}) < \varphi(\bar{x}) < \varphi(x_0).
\]
Consequently, one obtains
\[
\varphi(\hat{x}) - \varphi(x_0) = \varphi(\hat{x}) - \varphi(\bar{x}) + \varphi(\bar{x}) - \varphi(x_0) < \varphi(\bar{x}) - \varphi(x_0) = \tau.
\]
Because \( (\varphi(\hat{x}) - \varphi(x_0), g(\hat{x})) \) belongs to \( \mathcal{I}_{P,x_0}(S) \cap Q \), the last inequality contradicts the definition of \( \tau \). This completes the proof. \( \square \)

5.2. A zero-order Lagrangian optimality condition. Consider the standard mathematical programming problem with finitely many inequality and equality constraints, i.e.
\[
(\bar{P}) \quad \min_{x \in S} \varphi(x) \quad \text{subject to} \quad g_j(x) \leq 0, \quad j = 1, \ldots, m, \quad g_j(x) = 0, \quad j = m + 1, \ldots, p,
\]
where \( S \subseteq \mathbb{R}^n \), \( \varphi, g_j : \mathbb{R}^n \rightarrow \mathbb{R} \), \( j = 1, \ldots, p \) are given data. In what follows, its feasible region
\[
\bar{R} = \{ x \in S : \ g_j(x) \leq 0, \quad j = 1, \ldots, m, \quad g_j(x) = 0, \quad j = m + 1, \ldots, p \}
\]
will be assumed to be nonempty. Even if \( S \) happens to be strongly convex, the fact that problem data \( \varphi \) and \( g_j \) are merely \( C^{1,1}(\Omega) \), where \( \Omega \) is an open subset of \( \mathbb{R}^n \) containing \( S \), very often does not allow one to classify (\( \bar{P} \)) within convex optimization. This occurs already in quadratic programming and even if all \( g_j \), with \( j = 1, \ldots, p \), are convex functions, because of the presence of equality constraints. Nevertheless, under proper assumptions linking the radius of \( S \) with the regularity of the problem data, problem (\( \bar{P} \)) turns out to exhibit a convex geometry. Such an unexpected feature can be exploited in formulating Lagrangian type optimality
conditions. To see this, let us associate with \( \tilde{P} \) and with an element \( \bar{x} \in \tilde{R} \) the mapping \( I_{\tilde{P}, \bar{x}} : \mathbb{R}^n \rightarrow \mathbb{R}^{1+p} \) defined by
\[
I_{\tilde{P}, \bar{x}}(x) = (\varphi(x) - \varphi(\bar{x}), g_1(x), \ldots, g_p(x)).
\]

**Remark 5.2.** Notice that \( \tilde{P} \) is a special case of \( P \). Whenever \( \varphi, g_j \in C^{1,1}(\Omega) \) and \( S \) is strongly convex, problem \( \tilde{P} \) admits global solutions. A global solution \( \bar{x} \) to \( \tilde{P} \) is characterized by the condition
\[
I_{\tilde{P}, \bar{x}}(S) \cap ((-\infty, 0) \times (-\mathbb{R}^m_+ \times \{0\})) = \emptyset.
\]

The following optimality condition involves the classic Lagrangian function associated with problem \( \tilde{P} \), namely function \( L : \mathbb{R}^n \times \mathbb{R}^{1+p} \rightarrow \mathbb{R} \) defined by
\[
L(x, \lambda) = \lambda_0 \varphi(x) + \sum_{j=1}^p \lambda_j g_j(x).
\]

**Theorem 5.3.** With reference to problem \( \tilde{P} \), suppose that:
(i) \( S \) is \( r \)-convex, for some \( r > 0 \);
(ii) \( \varphi, g_j \in C^{1,1}(\Omega) \), with \( j = 1, \ldots, p \);
(iii) the following conditions are satisfied
\[
b_\tilde{P}^* > \text{lip}_{\tilde{P}} \quad \text{and} \quad r < \frac{b_\tilde{P}^*}{\text{lip}_{\tilde{P}}} - 1;
\]
If \( \bar{x} \in \tilde{R} \) is a global solution to \( \tilde{P} \), then there exists \( \bar{\lambda} = (\bar{\lambda}_0, \ldots, \bar{\lambda}_p) \in \mathbb{R}^{1+p} \backslash \{0\} \) such that
\[
\bar{\lambda}_j \geq 0, \quad j = 0, \ldots, m,
\]
\[
\bar{\lambda}_j g_j(\bar{x}) = 0, \quad j = 1, \ldots, m,
\]
and
\[
L(\bar{x}, \bar{\lambda}) = \min_{x \in S} L(x, \lambda).
\]
If, in particular, (iv) setting \( g = (g_1, \ldots, g_p) \), one has
\[
\bar{x} \in \tilde{R} \cap \text{int} S \quad \text{and} \quad b^*(Dg(\bar{x})) > 0,
\]
then \( \tilde{5.6} \) holds with \( \bar{\lambda}_0 = 1 \). Conversely, if conditions \( \tilde{5.4}, \tilde{5.5} \) and \( \tilde{5.6} \) hold true with \( \bar{\lambda}_0 > 0 \), then \( \bar{x} \in \tilde{R} \) is a global solution to \( \tilde{P} \).

**Proof.** The sufficiency part of the above optimality condition set can be obtained in an obvious manner and it is a well-known fact in mathematical programming.

Let us prove the necessary part. If \( \bar{x} \in \tilde{R} \) is a global minimizer for \( \tilde{P} \) the condition in \( \tilde{5.3} \) must be in force. Hypotheses (i), (ii) and (iii) enables one to apply Theorem 4.3, according to which \( I_{\tilde{P}, \bar{x}}(S) \) is a convex subset of \( \mathbb{R}^{1+p} \). If this is the situation, the geometric form of the Hahn-Banach theorem ensures the existence of a hyperplane of \( \mathbb{R}^{1+p} \) separating the two sets \( I_{\tilde{P}, \bar{x}}(S) \) and \((-\infty, 0] \times (-\mathbb{R}^m_+ \times \{0\}) \). In other terms, there exist a vector \( \lambda = (\lambda_0, \ldots, \lambda_p) \in \mathbb{R}^{1+p} \backslash \{0\} \) and \( \alpha \in \mathbb{R} \) such that
\[
(\lambda|w) \leq \alpha, \quad \forall w \in (-\infty, 0] \times (-\mathbb{R}^m_+ \times \{0\})
\]
and
\[
(\lambda|w) \geq \alpha, \quad \forall w \in I_{\tilde{P}, \bar{x}}(S),
\]
where \((\cdot|\cdot)\) denotes the inner product of an Euclidean space. As it is true that \( I_{\tilde{P}, \bar{x}}(\bar{x}) = (0, g_1(\bar{x}), \ldots, g_p(\bar{x})) \in I_{\tilde{P}, \bar{x}}(S) \cap ((-\infty, 0] \times (-\mathbb{R}^m_+ \times \{0\})) \), one obtains
\[
\sum_{j=1}^p \bar{\lambda}_j g_j(\bar{x}) = \alpha.
\]
Now, if it were $\alpha > 0$, since it is $2\mathcal{I}_{\tilde{P}, \tilde{x}}(\tilde{x}) \in (-\infty, 0] \times (-\mathbb{R}_+^m) \times \{0\}$, one would find by inequality (5.7)

$$
\sum_{j=1}^{p} \lambda_j g_j(\tilde{x}) \leq \frac{\alpha}{2},
$$

which is inconsistent with equality (5.9). Analogously, if it were $\alpha < 0$, since it is $\frac{1}{2} \mathcal{I}_{\tilde{P}, \tilde{x}}(\tilde{x}) \in (-\infty, 0] \times (-\mathbb{R}_+^m) \times \{0\}$, one would get again by inequality (5.7)

$$
\sum_{j=1}^{p} \lambda_j g_j(\tilde{x}) \leq 2\alpha,
$$

which is again inconsistent with equality (5.9). It remains to conclude that $\alpha = 0$, so that (5.9) becomes

(5.10)

$$
\sum_{j=1}^{p} \lambda_j g_j(\tilde{x}) = 0.
$$

Standard arguments, relying on a proper choice of elements in $(-\infty, 0] \times (-\mathbb{R}_+^m) \times \{0\}$, allow one to obtain from the last equality conditions (5.10) and (5.6). By using inequality (5.8), one obtains

$$
\lambda_0(\varphi(x) - \varphi(\bar{x})) + \sum_{j=1}^{p} \lambda_j g_j(x) \geq 0, \quad \forall x \in S,
$$

whence, by taking into account equality (5.10), conditions (5.6) immediately follows.

Now, suppose the additional hypothesis (iv) to hold and assume, ab absurdo, that $\lambda_0 = 0$. The Lyusternik-Graves theorem guarantees the linear openness of $g$ around the point $\tilde{x}$. This implies that there exist positive $\eta$ and $\zeta$ such that

$$
g(\text{int } B(\bar{x}, r)) \supseteq \text{int } B(g(\bar{x}), \eta r), \quad \forall r \in (0, \zeta].
$$

Since it is $\tilde{x} \in \text{int } S$, for a proper $r_0 \in (0, \zeta)$, one finds

$$
g(\tilde{x}) + \eta r_0 u \in g(\text{int } B(\bar{x}, r_0)), \quad \forall u \in B \subseteq \mathbb{R}^n,
$$

with $B(\bar{x}, r_0) \subseteq S$. Thus, by recalling inequality (5.8), on account of equality (5.10) one obtains

$$
((\bar{\lambda}_1, \ldots, \bar{\lambda}_p)|g(\tilde{x}) + \eta r_0 u) = \eta r_0 ((\bar{\lambda}_1, \ldots, \bar{\lambda}_p), u) \geq 0, \quad \forall u \in B,
$$

which leads to an absurdum. Consequently, $\lambda_0$ must be positive. By rescaling the value of $\lambda_j$, with $j = 1, \ldots, p$, condition (5.6) is proved to be satisfied as required in the last part of the thesis. Thus, the proof is complete. \hfill \Box

**Remark 5.4.** It should be clear that, since in the present context $Dg(\tilde{x})$ is represented by the Jacobian matrix of mapping $g$, calculated at $\tilde{x}$, the second condition in (iv) actually amounts to a full rank requirement.

### References

1. A.V. Arutyunov, *Covering mappings in metric spaces, and fixed points*, Dokl. Math. **76** (2007), no. 2, 665–668.
2. A.V. Arutyunov, *Stability of coincidence points and properties of covering mappings*, Math. Notes **86** (2009), no. 1-2, 153–158.
3. 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), no. 1, 105–127.
4. M.V. Balashov and D. Repovš, *On the splitting problem for selections*, J. Math. Anal. Appl. **355** (2009), no. 1, 277–287.
5. 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.
6. N.A. Bobylev, S.V. Emelyanov, and S.K. Korovin, *Convexity of images of convex sets under smooth maps*, Comput. Math. Model. **15** (2004), no. 3, 213–222.
7. J.M. Borwein and Q.J. Zhu, *Techniques of variational analysis*, Springer-Verlag, New York, 2005.
8. J.A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (1936), no. 3, 396–414.
9. A.V. Dmitruk, A.A. Milyutin and N.P. Osmolovskii, *Lyusternik’s theorem and the theory of extrema*, Russian Math. Surveys **35** (1980), no. 6, 11–51.
10. A.L. Dontchev and H. Frankowska, *Lyusternik-Graves Theorem and Fixed Points II*, J. Convex Anal. **19** (2012), no. 4, 955–973.
ON A CLASS OF CONVEX SETS WITH CONVEX IMAGES AND ITS APPLICATION TO NONCONVEX OPTIMIZATION

11. A.L. Dontchev and R.T. Rockafellar, *Implicit functions and solution mappings: A view from variational analysis*, Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

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. H. Frankowska and C. Olech, *R-convexity of the integral of set-valued functions*, Contributions to analysis and geometry (Baltimore, Md., 1980), pp. 117–129, Johns Hopkins Univ. Press, Baltimore, Md., 1981.

14. Giannessi, F., *Constrained optimization and image space analysis. Vol. 1. Separation of sets and optimality conditions*, Mathematical Concepts and Methods in Science and Engineering, 49. Springer, New York, 2005.

15. M.R. Hestenes, *Calculus of variations and optimal control theory*, John Wiley & Sons, Inc., New York-London-Sydney 1966.

16. A.D. Ioffe, *Regularity on a fixed set*, SIAM J. Optim. 21 (2011), no. 4, 1345–1370.

17. V.D. Milman, *Geometric theory of Banach spaces. II. Geometry of the unit ball*, Uspekhi Mat. Nauk 26 (1971), no. 6, 73–149 [in Russian].

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

19. B.T. Polyak, *The convexity principle and its applications*, Bull. Bras. Math. Soc. (N.S.) 34 (1) (2003) 59–75.

20. G. Reißig, *Computing abstractions of nonlinear systems*, IEEE Trans. Automat. Control 56 (2011), no. 11, 2583–2598.

21. R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

Dipartimento di Matematica e Applicazioni, Università di Milano-Bicocca, Via Cozzi, 55 - 20125 Milano, Italy

E-mail address: amos.uderzo@unimib.it