About Extensions of the Extremal Principle

Hoa T. Bui[1] · **Alexander Y. Kruger**[1]

Abstract In this paper, after recalling and discussing the conventional extremality, local extremality, stationarity and approximate stationarity properties of collections of sets and the corresponding (extended) extremal principle, we focus on extensions of these properties and the corresponding dual conditions with the goal to refine the main arguments used in this type of results, clarify the relationships between different extensions and expand the applicability of the generalized separation results. We introduce and study new more universal concepts of *relative extremality* and *stationarity* and formulate the *relative extended extremal principle*. Among other things, certain stability of the relative approximate stationarity is proved. Some links are established between the relative extremality and stationarity properties of collections of sets and (the absence of) certain regularity, lower semicontinuity and Lipschitz-like properties of set-valued mappings.

Dedicated to Professor Michel Théra on the occasion of his 70th birthday.

✉ Alexander Y. Kruger
a.kruger@federation.edu.au

Hoa T. Bui
h.bui@federation.edu.au

[1] Centre for Informatics and Applied Optimization, Faculty of Science and Technology, Federation University Australia, POB 663, Ballarat, Vic, 3350, Australia

1 Introduction

Starting with the pioneering work by Dubovitskii and Milyutin [6], it has become natural, when dealing with optimization and other related problems, to reformulate optimality and other properties under investigation as a kind of extremal behavior of certain collections of sets. The concept of a finite *extremal collection of sets* (Definition 1) represents a very general model embracing many optimality notions. It was first introduced in [27] and studied in [2, 9, 13, 15, 19, 26, 27, 30, 31] and many other publications. The *extremal principle* (Theorem 1) providing a dual characterization of (local) extremality in the form of *generalized separation* dates back to the 1980 paper by Kruger and Mordukhovich [27]. It can be considered as a generalization of the convex *separation theorem* to nonconvex sets and serves as a powerful tool for deducing optimality conditions in nonconvex optimization and subdifferential and coderivative calculus rules; cf. [15, 19, 30].

Similar to the classical analysis, besides extremality, the concepts of *stationarity* and *regularity* of collections of sets play an important role in this type of analysis and have been intensively investigated in recent years. It was established in [19, 20] that the conclusion of the extremal principle (the generalized separation) actually characterizes a much weaker than local extremality property of *approximate stationarity* (Definition 2). It became possible, keeping basically the original proof, to formulate (in the Asplund space setting) the *extended extremal principle*: the generalized separation is equivalent to the approximate stationarity (Theorem 2). The negation of the approximate stationarity happens to be an important regularity/transversality property of collections of sets known under various names [21–23, 25, 28] (A table illustrating the evolution of the terminology can be found in [25, Section 2].) and closely connected with the fundamental property of *metric regularity* of set-valued mappings.

Motivated by applications, there have been two independent attempts recently to single out the core part of the conventional proof of the extremal principle and formulate it as a separate statement with the objective to produce a more universal tool, applicable in situations where the conventional (extended) extremal principle fails: [24, Theorem 3.1] by Kruger and López and [35, Lemmas 2.1 and 2.2] by Zheng and Ng. The first result served as a tool when extending the extremal principle to infinite collections of sets, while the lemmas from [35] were used when proving fuzzy multiplier rules in set-valued optimization problems. These lemmas have been further refined and strengthened in Zheng and Ng [36, Theorems 3.1 and 3.4] and Zheng et al. [37, Theorem 1.1]. The mentioned statements in [35–37], in particular, waive the traditional for the extremal principle and its extensions in [19, 20, 24] assumption that the sets have nonempty intersection. Moreover, it has been observed in [36] that the conventional generalized separation condition can be strengthened by adding an additional condition (see condition (16)) determining the 'directions' of the dual space vectors. This additional condition is important, for instance, when recapturing the classical separation theorems.

In the current paper we study arbitrary (not necessarily smooth or convex) sets in a normed linear space. After recalling and discussing the conventional extremality, local extremality, stationarity and approximate stationarity properties of collections

of sets and the corresponding (extended) extremal principle, we focus on extensions of these properties and the corresponding dual conditions. The existing and some new extensions are considered with the goal to refine the main arguments used in this type of results, clarify the relationships between different extensions and expand the applicability of the generalized separation results.

We compare in detail the assumptions and conclusions in [24, Theorem 3.1] and [36, Theorem 3.4] and show (Proposition 5) that an appropriate reformulation of the latter theorem is a consequence of the first one. We also show (Corollary 2) that the main assertions in [37, Theorem 1.1] are consequences of the conventional extremal principle. At the same time, we demonstrate (Corollary 4) that [36, Theorem 3.4] is strong enough to recapture the nonlocal extremal principle, although it does not seem to be able to recapture the full local extremal principle. We briefly discuss (Remark 12.3) the role of the additional condition (16) and observe that it comes from subdifferentiating a norm at a nonzero point, and analogues of this condition are implicitly present in the proofs of the conventional extremal principle and its extensions. We admit the importance of conditions of the (16) type in generalized separation statements, but in the current paper, keeping in line with the conventional formulations and for the sake of simplicity of the presentation, we avoid adding such conditions to the statements.

Unlike the conventional extremal principle and its extensions in [19, 20, 24] assuming that the sets have a common point, in [35–37] the intersection of the sets is assumed to be empty, and each set is considered near its own point. This seems to be an important advancement, which in fact exploits the original ideas behind the conventional extremal principle. We demonstrate that the case of sets with empty intersection can still be treated within the conventional framework. This new point of view on the extremal principle is made explicit and further developed in the current paper introducing and studying the new more universal concepts of *relative extremality* and *stationarity*. We formulate the *relative extended extremal principle* (Theorem 5) and a 'relative' version of [24, Theorem 3.1] (Theorem 6). Among other things, certain stability of the relative approximate stationarity is proved (Proposition 13). Some links are established between the relative extremality and stationarity properties of collections of sets and (the absence of) certain regularity, lower semicontinuity and Lipschitz-like properties of set-valued mappings (Proposition 14 and Remark 16). As a consequence, we demonstrate a connection between the extremality and stationarity properties of collections of sets and the *nonconvex separation property* by Borwein and Jofre [2] (Proposition 15).

The structure of the paper is as follows. The next Section 2 contains some preliminary facts used throughout the paper. In Section 3 we recall and discuss the conventional definitions of extremality, local extremality, stationarity and approximate stationarity of pairs of sets, the conventional extremal principle and its extensions. The section contains some comparisons, illustrative examples and detailed historical comments. Section 4 is devoted to further extensions of the extremal principle. It contains a comparison of the assumptions and conclusions in [24, Theorem 3.1] and [36, Theorem 3.4], a study of the new concepts of extremality and stationarity relative to given points in each of the sets, the *relative extended extremal principle*, a 'relative' version of [24, Theorem 3.1], and a discussion of the links between the rela-

tive extremality and stationarity properties of collections of sets and (the absence of) certain regularity, lower semicontinuity and Lipschitz-like properties of set-valued mappings.

For simplicity, throughout the paper, we stick to the case of two nonempty sets, the general case of n ($n > 1$) sets not being strongly different. When formulating dual conditions (the extremal principle and its extensions), again for simplicity, only the Asplund space setting is considered. Recall that a Banach space is *Asplund* if every continuous convex function defined on an open convex set D is Fréchet differentiable at each point of some dense subset of D [32], or equivalently, if the dual of each its separable subspace is separable. We refer the reader to [3,30,32] for discussions about and characterizations of Asplund spaces. All reflexive, in particular, all finite dimensional Banach spaces are Asplund. By now it is well understood that extensions of the main results to broader classes of (or general) Banach spaces only require substituting in the proofs the Fréchet subdifferential sum rule with a sum rule for appropriate subdifferentials valid in such spaces. For instance, in general Banach spaces one can use Clarke subdifferentials or the classical convex subdifferentials if the sets are convex. One can also define certain abstract subdifferentials formulating the needed properties as axioms; see e.g. [24]. These are purely straightforward technical tricks which do not involve essentially new ideas.

2 Preliminaries

Our basic notation is standard, see e.g. [5,30,33]. Throughout the paper, X is a normed linear space. Its topological dual is denoted by X^* while $\langle \cdot, \cdot \rangle$ denotes the bilinear form defining the pairing between the two spaces. The closed unit balls in X and X^* are denoted by $\mathbb{B}$ and $\mathbb{B}^*$, respectively. $\mathbb{B}_\delta(x)$ denotes the open ball with radius $\delta > 0$ and center x. We use the same symbol $\| \cdot \|$ to denote norms in all normed linear spaces (primal and dual). If not explicitly stated otherwise, products of normed linear spaces are assumed to be equipped with the maximum norm: $\|(x,y)\| = \max\{\|x\|, \|y\|\}$, $(x,y) \in X \times Y$. For brevity, we sometimes write $\|x,y\|$ and $\mathbb{B}_\delta(x,y)$ instead of $\|(x,y)\|$ and $\mathbb{B}_\delta((x,y))$, respectively. Given a nonempty subset A of a normed linear space, int A and bd A stand, respectively, for its interior and boundary; $d(x,A) := \inf_{a \in A} \|x - a\|$ is the distance from a point x to A. We use the notation $\{A,B\}$ when referring to the pair of sets A and B as a single object. $\mathbb{N}$ stands for the set of all positive integers.

A set-valued mapping $F : X \rightrightarrows Y$ between two sets X and Y is a mapping, which assigns to every $x \in X$ a subset (possibly empty) $F(x)$ of Y. We use the notations gph $F := \{(x,y) \in X \times Y \mid y \in F(x)\}$ and dom $F := \{x \in X \mid F(x) \neq \emptyset\}$ for the graph and the domain of F, respectively, and $F^{-1} : Y \rightrightarrows X$ for the inverse of F. This inverse (which always exists with possibly empty values at some y) is defined by $F^{-1}(y) := \{x \in X \mid y \in F(x)\}$, $y \in Y$. Obviously, dom $F^{-1} = F(X)$.

Dual characterizations of extremality/stationarity (generalized separation) are formulated in this paper in terms of dual tools—Fréchet normal cones. Recall [19] that, given a subset A of a normed linear space X and a point $a \in A$, the *Fréchet normal*

cone to A at a is defined as follows:

$$N_A(a) := \left\{ x^* \in X^* \mid \limsup_{x \to a, x \in A \setminus \{a\}} \frac{\langle x^*, x - a \rangle}{\|x - a\|} \leq 0 \right\}. \tag{1}$$

It is a nonempty closed convex cone, often trivial ($N_A(\bar{x}) = \{0\}$). If A is a convex set, then (1) reduces to the normal cone in the sense of convex analysis:

$$N_A(a) := \{ x^* \in X^* \mid \langle x^*, x - a \rangle \leq 0 \quad \text{for all} \ x \in A \}.$$

Similarly, given a function $f : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ and a point $a \in A$ with $f(a) < \infty$, the closed convex set

$$\partial f(a) = \left\{ x^* \in X^* \mid \liminf_{x \to a, \ x \neq a} \frac{f(x) - f(a) - \langle x^*, x - a \rangle}{\|x - a\|} \geq 0 \right\}$$

is the *Fréchet subdifferential* of f at a. It reduces to the classical Moreau–Rockafellar subdifferential when f is convex. The following ε-extension ($\varepsilon \geq 0$) of (1) is used in the sequel: the set of ε-*normal elements* to A at $a \in A$:

$$N_\varepsilon(a \mid A) := \left\{ x^* \in X^* \mid \limsup_{x \to a, \ x \in A \setminus \{a\}} \frac{\langle x^*, x - a \rangle}{\|x - a\|} \leq \varepsilon \right\}. \tag{2}$$

When $\varepsilon = 0$, it reduces to (1). It is easy to check that $N_\varepsilon(a \mid A) \supset N_A(a) + \varepsilon \mathbb{B}$ for any $\varepsilon \geq 0$, and if A is not convex, the inclusion can be strict (see [14]).

     The following simple lemma used several times throughout the paper provides connections between two common ways of formulating 'generalized separation' in terms of normal cones. It is present implicitly in several existing proofs of dual conditions in the literature.

Lemma 1 *Let K_1 and K_2 be nonempty cones in a normed linear space and $\varepsilon \in (0, 1)$.*

(i) *Suppose vectors z_1 and z_2 satisfy the conditions:*

$$\|z_1\| + \|z_2\| = 1, \quad \|z_1 + z_2\| < \varepsilon, \quad z_1 \in K_1, \quad z_2 \in K_2.$$

Then there exist vectors $\hat{z}_1$ and $\hat{z}_2$ satisfying the following conditions:

$$\|\hat{z}_1\| + \|\hat{z}_2\| = 1, \quad \hat{z}_1 + \hat{z}_2 = 0, \quad d(\hat{z}_1, K_1) < \frac{\varepsilon}{2(1 - \varepsilon)}, \quad d(\hat{z}_2, K_2) < \frac{\varepsilon}{2(1 - \varepsilon)}.$$

(ii) *Suppose vectors z_1 and z_2 satisfy the conditions:*

$$\|z_1\| + \|z_2\| = 1, \quad z_1 + z_2 = 0, \quad d(z_1, K_1) + d(z_2, K_2) < \varepsilon.$$

Then there exist vectors $\hat{z}_1$ and $\hat{z}_2$ satisfying the following conditions:

$$\|\hat{z}_1\| + \|\hat{z}_2\| = 1, \quad \|\hat{z}_1 + \hat{z}_2\| < \frac{\varepsilon}{1 - \varepsilon}, \quad \hat{z}_1 \in K_1, \quad \hat{z}_2 \in K_2. \tag{3}$$

Proof (i) Set

$$z_1' := \frac{z_1 - z_2}{2}, \qquad z_2' := \frac{z_2 - z_1}{2}.$$

We have

$$z_1' + z_2' = 0, \quad \|z_1' - z_1\| = \|z_2' - z_2\| = \frac{1}{2}\|z_1 + z_2\| < \frac{\varepsilon}{2},$$

$$\|z_1'\| + \|z_2'\| \geq \|z_1\| + \|z_2\| - \|z_1' - z_1\| - \|z_2' - z_2\| > 1 - \varepsilon. \tag{4}$$

Now set

$$\hat{z}_1 := \frac{z_1'}{\|z_1'\| + \|z_2'\|}, \quad \hat{z}_2 := \frac{z_2'}{\|z_1'\| + \|z_2'\|}. \tag{5}$$

Then $\hat{z}_1 + \hat{z}_2 = 0$, $\|\hat{z}_1\| + \|\hat{z}_2\| = 1$ and, for $i = 1, 2$,

$$d(\hat{z}_i, K_i) \leq \left\| \hat{z}_i - \frac{z_i}{\|z_1'\| + \|z_2'\|} \right\| = \frac{\|z_i' - z_i\|}{\|z_1'\| + \|z_2'\|} < \frac{\varepsilon}{2(1 - \varepsilon)}.$$

(ii) There exist vectors $z_1' \in K_1$ and $z_2\prime \in K_2$ such that $\|z_1 - z_1'\| + \|z_2 - z_2'\| < \varepsilon$. Then condition (4) is satisfied,

$$\|z_1' + z_2'\| \leq \|z_1 + z_2\| + \|z_1 - z_1'\| + \|z_2 - z_2'\| < \varepsilon,$$

and the vectors $\hat{z}_1$ and $\hat{z}_2$ defined by (5) satisfy all the conditions in (3).   □

3 Extremality, Stationarity and Extremal Principle

In this section we recall and discuss the conventional definitions of extremality, local extremality, stationarity and approximate stationarity of pairs of sets, the conventional and extended extremal principles.

3.1 Extremality

Definition 1 (Extremality) Suppose X is a normed linear space, $A, B \subset X$ and $A \cap B \neq \emptyset$.

(i) The pair $\{A, B\}$ is *extremal* if for any $\varepsilon > 0$ there exist $u, v \in X$ such that

$$(A - u) \cap (B - v) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon; \tag{6}$$

(ii) The pair $\{A, B\}$ is *locally extremal* at $\bar{x} \in A \cap B$ if there exists a $\rho > 0$ such that for any $\varepsilon > 0$ there are $u, v \in X$ such that

$$(A - u) \cap (B - v) \cap \mathbb{B}_\rho(\bar{x}) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon. \tag{7}$$

Condition (i) (condition (ii)) in Definition 1 means that an appropriate arbitrarily small shift of the sets makes them nonintersecting (in a neighbourhood of $\bar{x}$). This is a very general model embracing many optimality notions. It is easy to see that, if a pair $\{A,B\}$ is extremal, it is locally extremal at any point in $A \cap B$, and the converse is true if A and B are convex. At the same time, the (nonlocal) extremality in condition (i) can be considered as a special case of the local extremality in condition (ii) with $\rho = \infty$.

The next example illustrates the difference between the extremality and the local extremality.

Example 1 1. The sets $A := \{(x_1, x_2) \mid x_2 \leq 0\}$ and $B := \{(x_1, x_2) \mid x_1^2 \leq x_2\}$ in $\mathbb{R}^2$ (see Fig. 1) are obviously extremal.

2. If the set A above is modified slightly: $A := \{(x_1, x_2) \mid x_2 \leq 0 \text{ or } x_1 \leq -1\}$ (see Fig. 2), then $\{A,B\}$ is not extremal any more. At the same time, it is still locally extremal at $(0,0) \in A \cap B$ (but not at $(-1,1)$!).
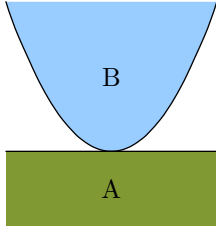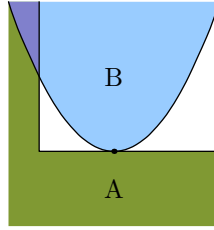

Fig. 1 Example 1.1   **Fig. 2** Example 1.2

As the next proposition shows, when the sets are closed, the first condition in (6) can be reformulated equivalently in terms of distances.

Proposition 1 (Distance characterizations of extremality) *Suppose X is a normed linear space, $A, B \subset X$ are closed and $A \cap B \neq \emptyset$. The pair $\{A,B\}$ is extremal if and only if for any $\varepsilon > 0$ there exist $u, v \in X$ such that $\max\{\|u\|, \|v\|\} < \varepsilon$ and the following two equivalent conditions hold:*

(i) $d(a - u, B - v) > 0$ *for all $a \in A$;*
(ii) $d(b - v, A - u) > 0$ *for all $b \in B$.*

Proof It is sufficient to show that each of the conditions (i) or (ii) is equivalent to $(A - u) \cap (B - v) = \emptyset$. Each of these conditions obviously implies $(A - u) \cap (B - v) = \emptyset$. Conversely, if $(A - u) \cap (B - v) = \emptyset$, then $a - u \notin B - v$ for any $a \in A$, and consequently, since B is closed, $d(a - u, B - v) > 0$, i.e., condition (i) is satisfied. Similarly, since A is closed, condition $(A - u) \cap (B - v) = \emptyset$ implies $d(b - v, A - u) > 0$ for all $b \in B$, hence, condition (ii). $\qquad\square$

The closedness assumption in Proposition 1 cannot be dropped.

Example 2 The pair of sets $A := \mathbb{R}^2 \setminus \{(t,0) \mid t > 0\}$ and $B := \{(0,0)\}$ in $\mathbb{R}^2$ is obviously extremal in the sense of Definition 1(i). At the same time, $d(a - u, B - v) = d(b - v, A - u) = 0$ for all $a \in A$, $b \in B$ and $u, v \in \mathbb{R}^2$.

Remark 1 Condition $(A - u) \cap (B - v) = \emptyset$ which is crucial for the extremality property in Definition 1 is obviously implied by the stronger condition $d(A - u, B - v) > 0$, which is also stronger than each of the conditions (i) or (ii) in Proposition 1. As the next example shows, condition $(A - u) \cap (B - v) = \emptyset$ in the definition of extremality cannot be replaced by condition $d(A - u, B - v) > 0$ even when both A and B are closed.

Example 3 Consider two sets in l^∞:

$$A := \left\{ x = (x^k) \mid x^K \in [K, K+1] \cup \left[K + 1 + \frac{1}{K}, K + 2 \right] \text{ for some } K \in \mathbb{N}; \right.$$
$$\left. x^k \in [-1, 1] \text{ for all } k \neq K \right\},$$
$$B := \left\{ x = (x^k) \mid x^K = K + 1 \text{ for some } K \in \mathbb{N}; x^k = 0 \text{ for all } k \neq K \right\}.$$

Observe that $A \cap B = B \neq \emptyset$. We are going to show that A and B are closed, $\{A, B\}$ is extremal, and $d(A - u, B - v) = 0$ for all $u, v \in l^\infty$ such that $\max\{\|u\|, \|v\|\} < 1/2$.

We first show that A is closed. Let $(x_n) \subset A$ and $x_n \to x_0 \in l^\infty$. There exist numbers $K, N \in \mathbb{N}$ such that $x_n^K \in [K, K+1] \cup [K + 1 + 1/K, K + 2]$ for all $n > N$. Indeed, assume on the contrary that for any $N > 0$ there exist $m, n > N$, $m \neq n$ and $K_m, K_n \in \mathbb{N}$, $K_n \neq K_m$ such that $x_m^{K_m} \in [K_m, K_m + 1] \cup [K_m + 1 + 1/K_m, K_m + 2]$ and $x_n^{K_n} \in [K_n, K_n + 1] \cup [K_n + 1 + 1/K_n, K_n + 2]$. Then

$$\|x_n - x_m\| \geq \max \left\{ |x_n^{K_n} - x_m^{K_n}|, |x_n^{K_m} - x_m^{K_m}| \right\}$$
$$\geq \max \left\{ |x_n^{K_n}| - |x_m^{K_n}|, |x_m^{K_m}| - |x_n^{K_m}| \right\} \geq \max\{K_n, K_m\} - 1 \geq 1,$$

which contradicts the assumption that (x_n) is convergent. Hence, $x_0 = (x_0^1, x_0^2, \ldots)$ with $x_0^K \in [K, K+1] \cup [K + 1 + 1/K, K + 2]$ and $x_0^k \in [-1, 1]$ for all $k \neq K$. Thus, A is closed.

A similar argument can be used to show that B is closed. Observe that for all $x_1, x_2 \in B$ with $x_1 \neq x_2$ one has $\|x_1 - x_2\| = \max\{K_1, K_2\} + 1$ for some $K_1, K_2 \in \mathbb{N}$, $K_1 \neq K_2$. Hence, $\|x_1 - x_2\| > 1$. It follows that any convergent sequence $(x_n) \subset B$ must be stationary when n is sufficiently large. This immediately yields the closedness of B.

Now we show that $\{A, B\}$ is extremal. Given an $\varepsilon \in (0, 1)$, find an $n \in \mathbb{N}$ such that $1/n < \varepsilon \leq 1/(n-1)$ and define a $u \in l^\infty$ as follows: $u^i = 1/n$ if $i < n$, and $u^i = 1/(i+1)$ if $i \geq n$. We have $\|u\| \leq \frac{1}{n} < \varepsilon$. Let $b = (b^k) \in B$, i.e., there exists a $K \in \mathbb{N}$ such that $b^K = K + 1$ and $b^k = 0$ for all $k \neq K$. Then $0 < (b+u)^k \leq 1/n < 1$ for all $k \neq K$. If $K < n$, then $(b+u)^K = K + 1 + 1/n$. If $K \geq n$, then $(b+u)^K = K + 1 + 1/(K+1)$. In both cases, $K + 1 < (b+u)^K < K + 1 + 1/K$. Hence, $b + u \notin A$, and consequently, $(A - u) \cap B = \emptyset$.

Let $u = (u^k), v = (v^k) \in l^\infty$ be such that $\max\{\|u\|, \|v\|\} < 1/2$ and $(A - u) \cap (B - v) = \emptyset$. We are going to show that $d(A - u, B - v) = 0$. Obviously $\|u - v\| < 1$.

Moreover, $|u^k - v^k| < 1/k$ for all $k \in \mathbb{N}$. Indeed, suppose on the contrary that $1/K \leq |u^K - v^K| < 1$ for some $K \in \mathbb{N}$ and choose a $b = (b^k) \in B$ such that $b^K = K + 1$ and $b^k = 0$ for $k \neq K$. Then for any $k \neq K$, we have $\|(b + u - v)^k\| = \|(u - v)^k\| < 1$, and $(b + u - v)^K = K + 1 + u^K - v^K$, and consequently, either $K < (b + u - v)^K < K + 1$ or $K + 1 + 1/K \leq (b + u - v)^K < K + 2$. In any case, $b + u - v \in A$, and $b - v \in A - u$, which is a contradiction. Thus, $|u^k - v^k| < 1/k$ for all $k \in \mathbb{N}$. For any $\varepsilon > 0$, we can find a $b \in B$ and a $K \in \mathbb{N}$ such that $b^K \neq 0$, $|u^K - u^K| < \varepsilon$. Set $a^K := b^K$ and $a^k := u^k - v^k$ for all $k \neq K$. Then $a = (a^k) \in A$ and $\|(a - u) - (b - v)\| = |a^K - b^K| < \varepsilon$. Hence, $d(A - u, B - v) = 0$.

3.2 Extremal Principle

The next well-known theorem (see Remark 3 below) gives approximate dual necessary conditions of local extremality in terms of Fréchet normals. It can be considered as a generalization of the classical convex *separation theorem* to pairs of nonconvex sets.

Theorem 1 (Extremal principle) *Suppose X is an Asplund space, $A, B \subset X$ are closed and $\bar{x} \in A \cap B$. If the pair $\{A, B\}$ is locally extremal at $\bar{x}$, then the following two equivalent conditions hold:*

(i) *for any $\varepsilon > 0$ there exist points $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$ and $a^* \in X^*$ such that*

$$\|a^*\| = 1, \quad d(a^*, N_A(a)) < \varepsilon \quad and \quad d(-a^*, N_B(b)) < \varepsilon; \tag{8}$$

(ii) *for any $\varepsilon > 0$ there exist points $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$, $a^* \in N_A(a)$ and $b^* \in N_B(b)$ such that*

$$\|a^*\| + \|b^*\| = 1 \quad and \quad \|a^* + b^*\| < \varepsilon. \tag{9}$$

Remark 2 The inequalities in (8) and (9) are only meaningful when $\varepsilon \leq 1$, because otherwise they are direct consequences of the corresponding equalities. Indeed, when $\varepsilon > 1$, condition (i) in Theorem 1 is satisfied automatically while condition (ii) guarantees only the existence of nontrivial normals in the ε-neighbourhood of $\bar{x}$ to at least one of the sets A and B, which is trivial as long as $\bar{x}$ is a boundary point of one of the sets (which is the case when $\{A, B\}$ is locally extremal at $\bar{x}$). If conditions (8) and (9) hold with $\varepsilon = 1$, they also hold with some $\varepsilon < 1$. Thanks to these observations, when applying Theorem 1 or its extensions, one can always assume that $\varepsilon < 1$.

Both conclusions in the above theorem are pretty common dual space properties used in many contemporary formulations of the extremal principle and its extensions. Properties (i) and (ii) can be found e.g. in, respectively, [30, Definition 2.5] (the *approximate extremal principle*) and [19, Definition 2.3] (the *generalized Euler equation*); cf. [23, property (SP)$_S$]. Condition (i) guarantees the existence of a pair of vectors a^* and b^* in the dual space, which are 'almost normal' (up to ε) to the corresponding sets at certain points with $a^* + b^* = 0$ and $\|a^*\| = \|b^*\| = 1$, while

condition (ii) guarantees the existence of a pair of vectors a^* and b^* which are exactly normal (in the Fréchet sense) to the corresponding sets at certain points with their sum $a^* + b^*$ being small (up to ε) and $\|a^*\| + \|b^*\| = 1$.

The equivalence of the two properties is a consequence of Lemma 1.

Proof of the equivalence of conditions (i) *and* (ii) *in Theorem 1*

(i) $\Rightarrow$ (ii). Take an arbitrary $\varepsilon > 0$ and set $\xi := \frac{\varepsilon}{1+\varepsilon}$. It follows from (i) that there exist points $a \in A \cap \mathbb{B}_\xi(\bar{x})$, $b \in B \cap \mathbb{B}_\xi(\bar{x})$ and $a^* \in X^*$ such that conditions (8) hold true with ξ in place of ε. Then

$$\left\| \frac{a^*}{2} \right\| + \left\| -\frac{a^*}{2} \right\| = 1 \quad \text{and} \quad d\left(\frac{a^*}{2}, N_A(a) \right) + d\left(-\frac{a^*}{2}, N_B(b) \right) < \xi.$$

Using Lemma 1(ii), we can find $\hat{a}^* \in N_A(a)$ and $\hat{b}^* \in N_B(b)$ such that

$$\|\hat{a}^*\| + \|\hat{b}^*\| = 1 \quad \text{and} \quad \|\hat{a}^* + \hat{b}^*\| < \frac{\xi}{1-\xi} = \varepsilon.$$

(ii) $\Rightarrow$ (i). Take an arbitrary $\varepsilon > 0$ and set $\xi := \frac{\varepsilon}{1+\varepsilon}$. It follows from (ii) that there exist points $a \in A \cap \mathbb{B}_\xi(\bar{x})$, $b \in B \cap \mathbb{B}_\xi(\bar{x})$, $a^* \in N_A(a)$ and $b^* \in N_B(b)$ such that conditions (9) hold true with ξ in place of ε. Using Lemma 1(i), we can find $\hat{a}^*, \hat{b}^* \in X^*$ such that

$$\|\hat{a}^*\| + \|\hat{b}^*\| = 1, \quad \hat{a}^* + \hat{b}^* = 0,$$
$$d(\hat{a}^*, N_A(a)) < \frac{\xi}{2(1-\xi)} = \frac{\varepsilon}{2}, \quad d(\hat{b}^*, N_B(b)) < \frac{\xi}{2(1-\xi)} = \frac{\varepsilon}{2}.$$

Then $\|2\hat{a}^*\| = 1$, $d(2\hat{a}^*, N_A(a)) < \varepsilon$ and $d(-2\hat{a}^*, N_B(b)) < \varepsilon$. $\square$

Conditions (i) and (ii) in Theorem 1 obviously hold if the pair $\{A,B\}$ is (not necessarily locally) extremal and $\bar{x} \in A \cap B$.

Remark 3 (Extremal principle: historical comments) The extremality properties in parts (i) and (ii) of Definition 1 were originally introduced in [27] (see Definition 4.1 and Remarks 4.1 and 4.8), where their connections with the separation of sets were also discussed and the first version of the extremal principle was established first in finite dimensions in terms of limiting normal cones [27, Theorem 4.1] and then extended, with the help of the Ekeland variational principle, to *Fréchet smooth* spaces, i.e., Banach spaces admitting an equivalent norm Fréchet differentiable away from zero [27, Theorem 6.1], in terms of sets of ε-normal elements. The latter result was formulated in the form similar to (but slightly weaker than) the condition (ii) in Theorem 1:

(ii)$'$ for any $\varepsilon > 0$ there exist points $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$, $a^* \in N_\varepsilon(a \mid A)$ and $b^* \in N_\varepsilon(b \mid B)$ such that conditions (9) hold true.

In the above condition, $N_\varepsilon(a \mid A)$ stands for the set of ε-normal elements (2) to A at a.

A slightly weaker version of [27, Theorem 6.1] (under the stronger assumption Definition 1(i) instead of (ii)) was presented in [26] accompanied by a short sketch of the proof.

While keeping the original pattern of the proof, the result of [27, Theorem 6.1] was strengthened in [15, Theorem 2] along two directions: 1) the assumption of the existence of an equivalent Fréchet differentiable norm was relaxed to that of the existence of a neighbourhood U of zero and a continuous function $\psi : U \to \mathbb{R}_+$ such that $\psi(x) = 0$ if and only if $x = 0$, ψ is Fréchet differentiable on $U \setminus \{0\}$ with $\|\psi(x)\| \geq 1$ for all $x \in U \setminus \{0\}$; and 2) conclusion (ii)$'$ replaced by a stronger one (the ε-*extremal principle* [30, Definition 2.5]):

(i)$'$ for any $\varepsilon > 0$ there exist points $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$ and $a^* \in X$ such that

$$\|a^*\| = 1, \quad a^* \in N_\varepsilon(a \,|\, A) \quad \text{and} \quad -a^* \in N_\varepsilon(b \,|\, B).$$

Note that this condition is still in general weaker than condition (i) in Theorem 1.

The next important step was made by Mordukhovich and Shao in [31, Theorem 3.2] where, using the subdifferential characterizations of Asplund spaces (the sum rule for Fréchet subdifferentials) established by Fabian [7, 8], the extremal principle with minimal adjustments in the original proof was extended to general Asplund spaces. In particular, it was shown that in Asplund spaces condition (i)$'$ above is equivalent to condition (i) in Theorem 1. Moreover, it was also shown in [31, Theorem 3.2] that Theorem 1 in its current form cannot be extended beyond Asplund spaces.

The last observation raised the question about possible extension of the extremal principle to non-Asplund spaces and the right tools needed for that since the Fréchet (ε-)normals cannot do the job. It did not take long for the experts in this area to pinpoint those properties of normals which are actually used in the conventional proof of the extremal principle. This led to several successful attempts to formulate these properties as sets of axioms and define several (very similar) abstract normal cones (and related subdifferentials) which could replace the Fréchet normal cones without changing much in the conventional proof of the extremal principle; see e.g. [2, 9] and [30, Subsection 2.5.3]. This allowed extending the extremal principle to *trustworthy* spaces (with respect to a given subdifferential/normal cone) [9] with Asplund spaces being trustworthy with respect to the Fréchet subdifferential and general Banach spaces being trustworthy with respect to, e.g., Clarke subdifferential.

Extending dual space results formulated in Asplund spaces in terms of Fréchet subdifferentials and normals, including the extremal principle, to general Banach spaces in terms of Clarke or other subdifferentials and normals, for which Banach spaces are trustworthy, has become a straightforward routine procedure. In this paper for simplicity we restrict the presentation to Asplund spaces and Fréchet normals only.

We refer the readers to [30, Section 2.6] for more historical comments.

Remark 4 (Extremal principle: two sets vs n sets) The original formulations of the definitions of extremality and extremal principle in [27] and most of their subsequent reformulations and generalizations [2, 3, 13, 15, 26, 30, 31] have been for the

more general than in Definition 1 and Theorem 1 setting of $n \geq 2$ sets. This seemingly more general setting is in fact not much different in terms of ideas, proofs and applications from the case of two sets considered for simplicity in the current paper. Moreover, it is well known (see e.g. [27, proof of Theorem 6.1], [13, p. 31], [15, proof of Theorem 2], [21, pp. 111 and 112], [30, proof of Theorem 2.10]) that the case of n sets $A_1, A_2, \ldots, A_n \subset X$ can be easily reduced to that of two sets: either $A := A_1 \times A_2 \times \cdots \times A_n$ and $B := \{(x, \ldots, x) \mid x \in X\}$ in X^n or $A := A_1 \times A_2 \times \cdots \times A_{n-1}$ and $B := \{(x, \ldots, x) \mid x \in A_n\}$ in X^{n-1}. This trick, sometimes referred to as *Pierra's product space reformulation* [34], is not easily applicable to the case of an infinite collection of sets treated in [24].

When dealing with arbitrary finite collections of sets, the seemingly weaker property of local extremality of a collection of n sets can be considered as a particular case of the nonlocal extremality of a collection of $n+1$ sets.

Remark 5 (Nonlocal extremality) The (nonlocal) extremality property, as defined in part (i) of Definition 1, does not use the assumption $A \cap B \neq \emptyset$, present in the preamble of Definition 1 as well as in the original definition of this property in [27]. The conventional proof of the extremal principle can proceed without this assumption (even getting a little shorter) and leading to a result (nonlocal extremal principle) which differs from the conclusions of Theorem 1 below by the conditions $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$ being replaced simply with $a \in A$, $b \in B$ and

$$\|a - b\| < d(A, B) + \varepsilon. \tag{10}$$

Note that if $A \cap B = \emptyset$, then $\{A, B\}$ is automatically extremal in the relaxed sense discussed in this remark.

Remark 6 (Normalization conditions) The two conditions (9) in Theorem 1 can be replaced by the following single one:

$$\|a^* + b^*\| < \varepsilon(\|a^*\| + \|b^*\|). \tag{11}$$

Observe that under condition (11) vectors a^* and b^* cannot equal zero simultaneously. The sum of the norms $\|a^*\| + \|b^*\|$ in (9) and (11) can be replaced by the maximum: $\max\{\|a^*\|, \|b^*\|\}$, or more generally, by $|||(\|a^*\|, \|b^*\|)|||_*$, where $|||\cdot|||_*$ is an arbitrary norm on $\mathbb{R}^2$. A similar observation can be made regarding the expression $\max\{\|u\|, \|v\|\}$ in Definition 1, where the maximum can be replaced by the sum: $\|u\| + \|v\|$, or more generally, by $|||(\|u\|, \|v\|)|||$, where $|||\cdot|||$ is an arbitrary norm on $\mathbb{R}^2$. In fact, it would be natural to choose the norms $|||\cdot|||$ and $|||\cdot|||_*$ to be dual to each other.

Theorem 1 formulated for a pair of sets yields the following result for a single set, generalizing (in Asplund spaces) the Bishop-Phelps theorem [32, Theorem 3.18] (cf. [13, Corollary 1 from Theorem 2.1], [31, Corollary 3.4], [19, Corollary 2.12.1], [30, Proposition 2.6]).

Corollary 1 (Density of 'support' points) *Suppose X is an Asplund space, $A \subset X$ is closed and $\bar{x} \in \text{bd}\,A$. Then, for any $\varepsilon > 0$, there exists a point $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$ such that $N_A(a) \neq \{0\}$.*

Proof The assertion follows by applying Theorem 1 to A and $B := \{\bar{x}\}$.   □

As observed in [30, p. 177], one can also go in the opposite direction: deducing a kind of extremal principle for a pair of sets from the density result for a single set in Corollary 1.

Proposition 2 *Suppose X is an Asplund space, $A, B \subset X$ and $A - B$ is closed (for instance, both sets are closed and one of them is compact). If the pair $\{A, B\}$ is extremal, then for any $\varepsilon > 0$ there exist points $a \in A$ and $b \in B$ satisfying $\|a - b\| < \varepsilon$, and an $a^* \in X^*$ such that*

$$\|a^*\| = 1, \quad a^* \in N_A(a) \quad and \quad -a^* \in N_B(b). \tag{12}$$

Proof Since $\{A, B\}$ is extremal, we have $0 \in A - B$ and $0 \notin \text{int}(A - B)$, hence, $0 \in \text{bd}(A - B)$. Given an $\varepsilon > 0$, by Corollary 1 there are $a \in A$ and $b \in B$ with $\|a - b\| < \varepsilon$, and an $a^* \in N_{A-B}(a - b)$ with $\|a^*\| = 1$. It remains to notice that the inclusion $a^* \in N_{A-B}(a - b)$ implies $a^* \in N_A(a)$ and $-a^* \in N_B(b)$ (see e.g. [19, Proposition 1.27]).   □

Remark 7 1. The conditions in (12) guaranteed by Proposition 2 are stronger than the corresponding conditions in (8) in Theorem 1. The latter conditions only guarantee that a^* and $-a^*$ are close to $N_A(a)$ and $N_B(b)$, respectively. On the other hand, unlike Theorem 1, Proposition 2 cannot relate the points $a \in A$ and $b \in B$ to a particular point in $A \cap B$.

2. The statement of Proposition 2 can be easily extended to the relaxed version of extremality without the assumption $A \cap B \neq \emptyset$ (see Remark 5). One only needs to replace the inequality $\|a - b\| < \varepsilon$ in the conclusion by condition (10). In the case $A \cap B = \emptyset$, as the next corollary shows, one can make another step and waive the assumption of the extremality of $\{A, B\}$.

Corollary 2 *Suppose X is an Asplund space, $A, B \subset X$ and $A - B$ is closed (for instance, both sets are closed and one of them is compact). If $A \cap B = \emptyset$, then for any $\varepsilon > 0$ there exist points $a \in A$ and $b \in B$ satisfying condition (10) and an $a^* \in X^*$ satisfying conditions (12).*

Proof Given an $\varepsilon > 0$, set $\varepsilon' = \varepsilon/2$. There exist points $a' \in A$ and $b' \in B$ satisfying $\|a' - b'\| < d(A, B) + \varepsilon'$. Without loss of generality $a' - b' \in \text{bd}(A - B)$. Indeed, since $A - B$ is closed and $0 \notin A - B$, the one-dimensional set $\{t \in [0, 1] \mid t(a' - b') \in A - B\}$ is compact and its infimum is attained at some $\bar{t} > 0$, which means that there exist $a'' \in A$ and $b'' \in B$ such that $a'' - b'' = \bar{t}(a' - b')$ and $a'' - b'' \in \text{bd}(A - B)$. Obviously $\|a'' - b''\| \leq \|a' - b'\| < d(A, B) + \varepsilon'$.

By Corollary 1, there are $a \in A$ and $b \in B$ with $\|(a - b) - (a' - b')\| < \varepsilon'$, and an $a^* \in N_{A-B}(a' - b')$ with $\|a^*\| = 1$. It remains to notice that $\|a - b\| < \|a' - b'\| + \varepsilon' < d(A, B) + \varepsilon$, and the inclusion $a^* \in N_{A-B}(a - b)$ implies $a^* \in N_A(a)$ and $-a^* \in N_B(b)$ (see e.g. [19, Proposition 1.27]).   □

Corollary 2 recaptures the main assertions in [37, Theorem 1.1].

3.3 Stationarity and Extended Extremal Principle

The extremal principle in Theorem 1 gives necessary conditions of (local) extremality which are in general not sufficient. Just like in the classical analysis and optimization theory, it actually characterizes a weaker than extremality property which can be interpreted as a kind of *stationarity*. The properties in the next proposition came to life as a result of a search for the weakest assumptions on the sets A and B which still ensure the conclusions of the extremal principle.

Definition 2 (Stationarity) Suppose X is a normed linear space, $A, B \subset X$ and $\bar{x} \in A \cap B$.

(i) The pair $\{A, B\}$ is *stationary* at $\bar{x}$ if for any $\varepsilon > 0$ there exist a $\rho \in (0, \varepsilon)$ and $u, v \in X$ such that

$$(A - u) \cap (B - v) \cap \mathbb{B}_\rho(\bar{x}) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon \rho; \qquad (13)$$

(ii) The pair $\{A, B\}$ is *approximately stationary* at $\bar{x}$ if for any $\varepsilon > 0$ there exist $\rho \in (0, \varepsilon)$, $a \in A \cap \mathbb{B}_\varepsilon(\bar{x})$, $b \in B \cap \mathbb{B}_\varepsilon(\bar{x})$ and $u, v \in X$ such that

$$(A - a - u) \cap (B - b - v) \cap (\rho \mathbb{B}) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon \rho. \qquad (14)$$

Unlike (7), in conditions (13) and (14) the size of the "shifts" of the sets is related to that of the neighborhood in which the sets become nonintersecting, namely $\max\{\|u\|, \|v\|\}/\rho < \varepsilon$. Compared to (13), in conditions (14), instead of the common point $\bar{x}$, the sets A and B are considered near their own points a and b, respectively.

The implications in the next proposition are immediate consequences of Definitions 1 and 2, while the equivalences were proved in [21, Proposition 14].

Proposition 3 (Extremality vs stationarity) *Suppose X is a normed linear space, $A, B \subset X$ and $\bar{x} \in A \cap B$. Consider the following properties:*

 (i) $\{A, B\}$ *is extremal;*
 (ii) $\{A, B\}$ *is locally extremal at $\bar{x}$;*
(iii) $\{A, B\}$ *is stationary at $\bar{x}$;*
(iv) $\{A, B\}$ *is approximately stationary at $\bar{x}$.*

Then (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). *If, additionally, A and B are convex, then* (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv).

All the implications in Proposition 3 can be strict; see some examples in [21, 23].

Replacing in Theorem 1 local extremality with approximate stationarity produces a stronger statement—the *extended extremal principle*, with the two equivalent conditions in the conclusion of Theorem 1 becoming not only necessary but also sufficient, thus producing full duality. The proof of the necessity in the next theorem is a refined version of the proof of Theorem 1, while the proof of the sufficiency is a straightforward consequence of the definitions and does not use the assumption of the Asplund property of the space; cf. [18, Theorem 4.1].

Theorem 2 (Extended extremal principle) *Suppose X is an Asplund space, $A, B \subset X$ are closed and $\bar{x} \in A \cap B$. The pair $\{A, B\}$ is approximately stationary at $\bar{x}$ if and only if the two equivalent conditions in Theorem 1 hold true.*

Remark 8 (Extended extremal principle: historical comments) The approximate stationarity property in part (ii) of Definition 2 was originally introduced in a slightly different form in [16, formula (4)], where the property was referred to as *extremality near $\bar{x}$*. A version of the extended extremal principle was formulated in [16, Theorem 2] in the setting of a Fréchet smooth Banach space in the form of condition (i)$'$ in Remark 3. The property in part (i) of Definition 2 was also implicitly present in [16] (see formula (5)). The result was extended to Asplund spaces in [17, Theorem 2], where it was formulated in the form of condition (ii)$'$ in Remark 3. The full proof of the extended extremal principle in the form of condition (ii) in Theorem 1 appeared in [18, Theorem 4.1], where the property in part (ii) of Definition 2 was referred to as *extended extremality (e-extremality) near $\bar{x}$*. It was shown in [19, Theorem 3.7] that in its current form the result cannot be extended beyond Asplund spaces.

In [20–22] the properties in Definition 2 are referred to as *stationarity* and *weak stationarity*, respectively. [22] gives a slightly improved version of the definition of the last property, compared to that in [20, 21]. The name *approximate stationarity* for the property in part (ii) of Definition 2 appeared in [23]. Extensions of Theorem 2 to non-Asplund spaces are discussed in [24].

Remark 9 (Extended extremal principle: two sets vs n sets) Similarly to the case of extremality and extremal principle (see Remark 4), the stationarity properties and extended extremal principle are usually formulated for the setting of $n \geq 2$ sets with the case of n sets easily reduced the same way to that of two sets (see e.g. [16, item 3], [17, Definition 4], [18, Definition 4.2], [19, Definition 3.5], [20, Proposition 7]), [21, Proposition 20 and Remark 6]). Remark 6 applies entirely to Definition 2 and Theorem 2.

Remark 10 (Approximate stationarity vs transversality) Theorem 2 can be reformulated as equivalence of the negations of the primal and dual properties involved in its statement: the absence of the approximate stationarity is equivalent to the absence of the generalized separation. These are important *regularity/transversality* properties of pairs of sets involved in constraint qualifications, qualification conditions in subdifferential calculus and convergence analysis of computational algorithms [11, 21–23, 28, 29]. They are known under various names. A table illustrating the evolution of the terminology can be found in [25, Section 2]. The transversality property of finite collections of sets (the negation of the approximate stationarity) is in a sense equivalent to the famous *metric regularity* property of set-valued mappings; cf. [21–23]. See also the discussion in Subsection 4.3 below.

4 More Extensions

4.1 Two Recent Extensions

As it was pointed out in Remark 3, the key tool used in the proof of Theorem 1 (and also Theorem 2) is the Ekeland variational principle. Theorem 2 is in a sense the ultimate version of Theorem 1 establishing the same conclusion (generalized separation) under the weakest possible assumptions on the pair of sets (approximate stationarity), thus, providing the complete duality (in the Asplund space setting) between the corresponding primal space and dual space properties.

The next natural step in the extremal principle refinement process is to single out the core part of the conventional proof of the extremal principle around the application of the Ekeland variational principle, identify the minimal assumptions on the sets and the immediate conclusions and formulate it as a separate statement. Such a result (results) would expose the core arguments behind the extremal principle and could serve as a key building block when constructing other generalized separation statements, applicable in situations where the conventional (extended) extremal principle fails.

We are aware of two recent attempts of this kind: [24, Theorem 3.1] which served as a tool when extending Theorems 1 and 2 to infinite collections of sets, and [35, Lemmas 2.1 and 2.2] used when proving fuzzy multiplier rules in set-valued optimization problems. The last couple of lemmas have been further refined and strengthened in [36, Theorems 3.1 and 3.4] and [37, Theorem 1.1].

The next two theorems are reformulations for the setting adopted in the current paper of [24, Theorem 3.1] and [36, Theorem 3.4], respectively.

Theorem 3 (Kruger and López, 2012) *Suppose X is an Asplund space, $A, B \subset X$ are closed and $\bar{x} \in A \cap B$.*

(i) *If points $a \in A$, $b \in B$ and $u, v \in X$ satisfy conditions* (14) *with some $\varepsilon > 0$ and $\rho > 0$, then, for any $\delta > \max\{\|a - \bar{x}\|, \|b - \bar{x}\|\} + \rho(\varepsilon + 1)$, there exist points $a' \in A \cap \mathbb{B}_\delta(\bar{x})$, $b' \in B \cap \mathbb{B}_\delta(\bar{x})$ and $a^* \in N_A(a')$, $b^* \in N_B(b')$ satisfying conditions* (9).

(ii) *If $a \in A$, $b \in B$ and $a^* \in N_A(a)$, $b^* \in N_B(b)$ satisfy conditions* (9) *for some $\varepsilon > 0$, then, for any $\delta > 0$, there exist a $\rho \in (0, \delta)$ and points $u, v \in X$ satisfying conditions* (14).

Theorem 4 (Zheng and Ng, 2011) *Suppose X is an Asplund space, $A, B \subset X$ are closed, $A \cap B = \emptyset$. If points $a \in A$ and $b \in B$ satisfy condition* (10) *with some $\varepsilon > 0$, then, for any $\lambda > 0$ and $\tau \in (0, 1)$, there exist points $a' \in A \cap \mathbb{B}_\lambda(a)$, $b' \in B \cap \mathbb{B}_\lambda(b)$ and $a^* \in X^*$ such that*

$$\|a^*\| = 1, \qquad d(a^*, N_A(a')) + d(-a^*, N_B(b')) < \varepsilon/\lambda, \tag{15}$$

$$\tau\|a' - b'\| \le \langle a^*, a' - b' \rangle. \tag{16}$$

First observe that the extremal principle in Theorem 2 is a direct corollary of Theorem 3.

Proof (Theorem 2 from Theorem 3) Let the pair $\{A,B\}$ be approximately stationary at $\bar{x}$. We are going to show that condition (ii) in Theorem 1 holds true. Given an $\varepsilon > 0$, find an $\varepsilon' > 0$ such that $\varepsilon'(\varepsilon' + 2) < \varepsilon$. By Definition 2(ii), there exist $\rho \in (0, \varepsilon')$, $a \in A \cap \mathbb{B}_{\varepsilon'}(\bar{x})$, $b \in B \cap \mathbb{B}_{\varepsilon'}(\bar{x})$ and $u, v \in X$ such that conditions (14) are satisfied with ε' in place of ε. Then $\max\{\|a - \bar{x}\|, \|b - \bar{x}\|\} + \rho(\varepsilon' + 1) < \varepsilon' + \varepsilon'(\varepsilon' + 1) < \varepsilon$, and it follows from Theorem 3(i) that there exist points $a' \in A \cap \mathbb{B}_{\varepsilon}(\bar{x})$, $b' \in B \cap \mathbb{B}_{\varepsilon}(\bar{x})$, $a^* \in N_A(a')$ and $b^* \in N_B(b')$ satisfying (9), i.e. condition (ii) in Theorem 1 holds true.

Conversely, let condition (ii) in Theorem 1 holds true and an $\varepsilon > 0$ be given. Then there exist points $a \in A \cap \mathbb{B}_{\varepsilon}(\bar{x})$, $b \in B \cap \mathbb{B}_{\varepsilon}(\bar{x})$, $a^* \in N_A(a)$ and $b^* \in N_B(b)$ satisfying conditions (9). By Theorem 3(ii), there exist a $\rho \in (0, \varepsilon)$ and points $u, v \in X$ satisfying conditions (14), i.e., the pair $\{A,B\}$ is approximately stationary at $\bar{x}$. □

Remark 11 The observations in Remark 2 are applicable to Theorem 3(i): we can always assume that $\varepsilon < 1$. Similarly, in Theorem 4 we can assume that $\varepsilon < \lambda$.

Next we compare the statements of Theorem 3(i) and Theorem 4. There are important similarities between them: both establish a kind of generalized separation of the two sets, related somehow to the given pair of points $a \in A$ and $b \in B$ possessing a certain approximate 'extremality' property. There are also essential differences.

We start with comparing the assumptions in the two statements. On the first glance, they look mutually exclusive: the first one assumes the existence of a point $\bar{x} \in A \cap B$, while in the second theorem, it is assumed on the contrary that $A \cap B = \emptyset$. However, this distinction is easy to overcome. Given points $a \in A$ and $b \in B$ in Theorem 4, one can set $A' := A - a$ and $B' := B - b$; then $\bar{x} := 0 \in A' \cap B'$ (this trick is used in the proof of Theorem 4' below). This observation exposes also the different roles played by the pairs $a \in A$ and $b \in B$ in Theorem 3(i) and Theorem 4. In the first one, these are actually additional parameters having no analogues in Theorem 4, which corresponds to $a = b = \bar{x}$ in Theorem 3(i).

The second distinction is related to the main approximate 'extremality' assumptions on the pair of sets: conditions (14) in Theorem 3(i) and condition (10) in Theorem 4. The next proposition shows that condition (10) implies a stronger version of conditions (14).

Proposition 4 (Conditions (14) vs condition (10)) *Suppose X is a normed linear space, $A, B \subset X$, $A \cap B = \emptyset$ and $\varepsilon > 0$. If points $a \in A$ and $b \in B$ satisfy condition (10), then there exist $u, v \in X$ such that $\|u\| = \|v\| < \frac{\varepsilon}{2}$ and*

$$(A - a - u) \cap (B - b - v) = \emptyset. \tag{17}$$

Moreover, one can take

$$u := \frac{\varepsilon'}{2} \cdot \frac{b - a}{\|b - a\|} \quad and \quad v := \frac{\varepsilon'}{2} \cdot \frac{a - b}{\|b - a\|}, \tag{18}$$

where, if $d(A, B) > 0$, then ε' can be any number satisfying $\|b - a\| - d(A, B) < \varepsilon' < \min\{\varepsilon, \|b - a\|\}$, and if $d(A, B) = 0$, then $\varepsilon' = \|b - a\| < \varepsilon$.

Proof Suppose u and v are given by (18) and condition (17) is violated. Then $\hat{a} - a - u = \hat{b} - b - v$ for some $\hat{a} \in A$ and $\hat{b} \in B$, and

$$\hat{b} - \hat{a} = b - a - (u - v) = \left(1 - \frac{\varepsilon'}{\|b - a\|}\right)(b - a). \tag{19}$$

If $d(A,B) > 0$, then $\varepsilon' < \|b - a\|$, and consequently,

$$\|b - a\| - d(A,B) \geq \|b - a\| - \|\hat{b} - \hat{a}\| = \|b - a\| - \left(1 - \frac{\varepsilon'}{\|b - a\|}\right)\|b - a\| = \varepsilon',$$

which contradicts the choice of ε'. If $d(A,B) = 0$, then $\varepsilon' = \|b - a\|$, and it follows from (19) that $\hat{b} = \hat{a}$ which contradicts the assumption that $A \cap B = \emptyset$. Thus, condition (17) is true.   □

Proposition 4 is not reversible: condition (17) being satisfied with some small u and v does not imply that $\|a - b\|$ is close to the distance $d(A,B)$ between the two sets.

Example 4 Let $A := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \leq 0\}$ and $B := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \geq 1\}$. Then, assuming that $\mathbb{R}^2$ is equipped with e.g. the sum norm, $d(A,B) = 1$. If $a := (\alpha,0) \in A$ and $b := (\beta,1) \in B$ with some $\alpha,\beta \in \mathbb{R}$, then condition (17) is satisfied with $u := (0,\varepsilon)$, $v := (0,-\varepsilon)$ and any $\varepsilon > 0$. At the same time, $\|a - b\| = |\alpha - \beta| + 1$ can be arbitrarily large when the numbers α and β are far apart.

Thus, condition (17) with small u and v is less restrictive than condition (10). Moreover, the first condition in (14) with $\rho < \infty$ is weaker than (17) and allows for local versions of the corresponding properties.

The next assertion is immediate from Proposition 4.

Corollary 3 *Suppose X is a normed linear space, $A,B \subset X$, $A \cap B = \emptyset$. If sequences $\{a_k\} \subset A$ and $\{b_k\} \subset B$ are such that $\|a_k - b_k\| \to d(A,B)$, then there exist sequences $\{u_k\},\{v_k\} \subset X$ converging to 0, such that*

$$(A - a_k - u_k) \cap (B - b_k - v_k) = \emptyset.$$

With the sets in Example 4, one can easily see that the statement of Corollary 3 is not reversible.

Now we are going to compare the conclusions of the two theorems. Similarly to the two conditions in Theorem 1, they represent two different ways of formulating dual extremality/separation conditions: in terms of normal (in Theorem 3(i)) or 'almost normal' (in Theorem 4) vectors, with the connection between the two formulations provided by Lemma 1. However, unlike the two equivalent conditions in Theorem 1 formulated 'for any $\varepsilon > 0$', in both Theorem 3(i) and Theorem 4 the number $\varepsilon > 0$ is a given quantitative parameter. Lemma 1 used in the proof of the equivalence of the two conditions in Theorem 1 cannot provide one-to-one translation between the two settings with the given $\varepsilon > 0$; it only gives estimates, and its application leads to some 'loss of accuracy'. Note that the proof of [24, Theorem 3.1], where Theorem 3 is taken from, contains estimates in terms of 'almost normal' vectors and then

employs the arguments used in the proof of Lemma 1(ii) to ensure that the vectors belong to the normal cones. To make a fair comparison, one needs to either reformulate Theorem 4 in terms of normal vectors using Lemma 1, or extract the pre-Lemma 1 statement from the proof of [24, Theorem 3.1]. Below for simplicity we follow the first approach. The next statement is a consequence of Theorem 4 and Lemma 1(ii).

Theorem 4′ *Suppose X is an Asplund space, $A, B \subset X$ are closed, $A \cap B = \emptyset$. If points $a \in A$ and $b \in B$ satisfy condition (10) with some $\varepsilon > 0$, then, for any $\lambda > 0$, there exist points $a' \in A \cap \mathbb{B}_{\frac{\varepsilon+\lambda}{2}}(a)$, $b' \in B \cap \mathbb{B}_{\frac{\varepsilon+\lambda}{2}}(b)$, $a^* \in N_A(a')$ and $b^* \in N_B(b')$ such that*

$$\|a^*\| + \|b^*\| = 1 \quad \text{and} \quad \|a^* + b^*\| < \varepsilon/\lambda. \tag{20}$$

Proof Given $\varepsilon > 0$ and $\lambda > 0$, set $\lambda' := \frac{\varepsilon+\lambda}{2}$. Thanks to Theorem 4, there exist points $a' \in A \cap \mathbb{B}_{\lambda'}(a), b' \in B \cap \mathbb{B}_{\lambda'}(b)$ and $a^* \in X^*$ satisfying conditions (15) with λ' in place of λ. Then

$$\left\|\frac{a^*}{2}\right\| + \left\|-\frac{a^*}{2}\right\| = 1 \quad \text{and} \quad d\left(\frac{a^*}{2}, N_A(a')\right) + d\left(-\frac{a^*}{2}, N_B(b')\right) < \frac{\varepsilon}{2\lambda}.$$

Using Lemma 1(ii), we can find $\hat{a}^* \in N_A(a')$ and $\hat{b}^* \in N_B(b')$ such that

$$\|\hat{a}^*\| + \|\hat{b}^*\| = 1 \quad \text{and} \quad \|\hat{a}^* + \hat{b}^*\| < \frac{\varepsilon/(2\lambda')}{1 - \varepsilon/(2\lambda')} = \frac{\varepsilon}{2\lambda' - \varepsilon} = \frac{\varepsilon}{\lambda}.$$

$\square$

Now the comparison is straightforward.

Proposition 5 *Theorem 3(i) implies Theorem 4′.*

Proof Under the conditions of Theorem 4′, set $A' := A - a$ and $B' := B - b$. Then $0 \in A' \cap B'$. By Proposition 4, there exist $u, v \in X$ such that $\|u\| = \|v\| < \frac{\varepsilon}{2}$ and $(A' - u) \cap (B' - v) = \emptyset$. Choose an $\varepsilon' \in (2\|u\|, \varepsilon)$ and set $\rho := \lambda/2$, $\hat{\varepsilon} := \varepsilon'/\lambda$ and $\delta := (\varepsilon + \lambda)/2$. Then $\hat{\varepsilon} < \varepsilon/\lambda$ and $\delta > \rho(\hat{\varepsilon} + 1)$. By Theorem 3(i) applied to the sets A' and B' at 0, there exist points $a' \in A \cap \mathbb{B}_\delta(a)$, $b' \in B \cap \mathbb{B}_\delta(b)$, $a^* \in N_A(a')$ and $b^* \in N_B(b')$ satisfying conditions (20). $\square$

Thus, Theorem 4′ is a special case of Theorem 3(i). On the other hand, as demonstrated in [35, 36], Theorem 4 (as well as its version formulated above as Theorem 4′) is sufficient for many important applications. Next we show that Theorem 4′ implies the nonlocal version of the extremal principle.

Corollary 4 (Nonlocal extremal principle) *Suppose X is an Asplund space, $A, B \subset X$ are closed and $\bar{x} \in A \cap B$. If the pair $\{A, B\}$ is extremal at $\bar{x}$, then the two equivalent conditions in Theorem 1 hold true.*

Proof Let the pair $\{A, B\}$ be extremal and a number $\varepsilon > 0$ be given. Choose an $\varepsilon' \in \left(0, \frac{\varepsilon^2}{\varepsilon+1}\right)$. Then $\frac{\varepsilon'}{\varepsilon} < \varepsilon - \varepsilon'$ and we can choose a λ such that $\frac{2\varepsilon'}{\varepsilon} < \lambda < 2(\varepsilon - \varepsilon')$. There exist vectors $u, v \in X$ satisfying conditions (6) with ε' in place of ε. Define $A' := A - u$, $B' := B - v$, $a := \bar{x} - u$ and $b := \bar{x} - v$. Then $A' \cap B' = \emptyset$ and $\|a - b\| < 2\varepsilon'$. Applying Theorem 4', we find points $a' \in A$, $b' \in B$, $a^* \in N_{A'}(a' - u) = N_A(a')$ and $b^* \in N_{B'}(b' - v) = N_B(b')$ such that $\max\{\|a' - \bar{x}\|, \|b' - \bar{x}\|\} < \varepsilon' + \lambda/2 < \varepsilon$, $\|a^*\| + \|b^*\| = 1$ and $\|a^* + b^*\| < 2\varepsilon'/\lambda < \varepsilon$. Thus, condition (ii) in Theorem 1 is satisfied. □

Remark 12 1. It is not difficult to modify the proof of Corollary 4 to cater for the relaxed version of nonlocal extremality without the assumption $A \cap B \neq \emptyset$ (see Remark 5).

2. Theorem 4 does not seem to be able to recapture the full local extremal principle (as in Theorem 1), not to say the extended extremal principle (as in Theorem 2).

3. Condition (16) in Theorem 4 determining the 'direction' of the vector a^* does not have a direct analogue in the statement of Theorem 3(i). Together with the first condition in (15), it comes from subdifferentiating a norm at a nonzero point in the proof of [36, Theorem 3.4]. Subdifferentiating a norm is an essential component also in the proofs of the conventional extremal principle and all its modifications, including the one in [24, Theorem 3.1]; so analogues of (16) are implicitly present in all such proofs. Zheng and Ng [36] seem to be the first to notice the importance of conditions like (16) for recapturing the classical convex separation theorem, and make (16) explicit in the statement of [36, Theorem 3.4]. In the current paper, keeping in line with the conventional formulations and for the sake of simplicity of the presentation, we will not formulate analogues of the condition (16) in the subsequent statements.

4.2 Relative Extremality and Stationarity

The conventional definition of extremality (Definition 1) and most of its extensions presume that the sets have a common point. However, as it was demonstrated in Subsection 4.1, there are natural situations which allow for and, in fact, require application of the extremal principle or its extensions to sets with empty intersection. Such situations are formalized in the current subsection.

As it was observed in Remark 5, the nonlocal extremality property in Definition 1(i) does not use the assumption $A \cap B \neq \emptyset$, and the conventional proof of the extremal principle can proceed without this assumption. Now we are going to relax the definitions of local extremality and stationarity properties of pairs of sets. Instead of considering both sets near a common point, we are going to consider each set near its own point. The next definition builds on the simple trick employed in the proof of Proposition 5 and present implicitly already in Definition 2(ii) of approximate stationarity and, in view of Proposition 9 below, even in Definition 1.

Definition 3 (Relative extremality and stationarity) Suppose X is a normed linear space, $A, B \subset X$, $a \in A$ and $b \in B$.

(i) The pair $\{A,B\}$ is *extremal* relative to a and b if the pair $\{A-a,B-b\}$ is extremal, i.e., for any $\varepsilon > 0$ there are $u,v \in X$ such that

$$(A-a-u) \cap (B-b-v) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon.$$

(ii) The pair $\{A,B\}$ is *locally extremal* relative to a and b if the pair $\{A-a,B-b\}$ is locally extremal at 0, i.e., there exists a $\rho > 0$ such that for any $\varepsilon > 0$ there are $u,v \in X$ such that

$$(A-a-u) \cap (B-b-v) \cap (\rho \mathbb{B}) = \emptyset \quad \text{and} \quad \max\{\|u\|, \|v\|\} < \varepsilon. \qquad (21)$$

(iii) The pair $\{A,B\}$ is *stationary* relative to a and b if the pair $\{A-a,B-b\}$ is stationary at 0, i.e., for any $\varepsilon > 0$ there exist a $\rho \in (0,\varepsilon)$ and $u,v \in X$ such that conditions (14) hold true.

(iv) The pair $\{A,B\}$ is *approximately stationary* relative to a and b if the pair $\{A-a,B-b\}$ is approximately stationary at 0, i.e., for any $\varepsilon > 0$ there exist a $\rho \in (0,\varepsilon)$ and points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ and $u,v \in X$ such that conditions (14) with a' and b' in place of a and b hold true.

Definition 3 reduces the extremality, local extremality, stationarity and approximate stationarity at individual points to the corresponding conventional properties in the sense of Definitions 1 and 2. On the other hand, Definitions 1 and 2 are special cases of the corresponding items in Definition 3 when $a = b = \bar{x}$. From Proposition 3 we get the following statement.

Proposition 6 (Extremality vs stationarity) *Suppose X is a normed linear space, $A,B \subset X$, $a \in A$ and $b \in B$. Consider the following properties:*

(i) *$\{A,B\}$ is extremal relative to a and b;*
(ii) *$\{A,B\}$ is locally extremal relative to a and b;*
(iii) *$\{A,B\}$ is stationary relative to a and b;*
(iv) *$\{A,B\}$ is approximately stationary relative to a and b.*

Then (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). *If, additionally, A and B are convex, then* (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv).

Similarly, the next two theorems generalizing the conventional extended extremal principle in Theorem 2 and its extension in Theorem 3 to the case of individual points are direct corollaries of Theorems 2 and 3, respectively.

Theorem 5 (Relative extended extremal principle) *Suppose X is an Asplund space, $A,B \subset X$ are closed, $a \in A$ and $b \in B$. The pair $\{A,B\}$ is approximately stationary relative to a and b if and only if the following two equivalent conditions hold:*

(i) *for any $\varepsilon > 0$ there exist points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ and $a^* \in X^*$ satisfying conditions (8) with a' and b' in place of a and b, respectively;*
(ii) *for any $\varepsilon > 0$ there exist points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$, $a^* \in N_A(a')$ and $b^* \in N_B(b')$ satisfying conditions (9).*

Theorem 6 ('Relative' version of Theorem 3) *Suppose X is an Asplund space, $A,B \subset X$ are closed, $a \in A$ and $b \in B$.*

(i) *If points $\hat{a} \in A$, $\hat{b} \in B$ and $u, v \in X$ satisfy conditions (14) with $\hat{a}$ and $\hat{b}$ in place of a and b, respectively, and some $\varepsilon > 0$ and $\rho > 0$, then, for any $\delta > \max\{\|\hat{a} - a\|, \|\hat{b} - b\|\} + \rho(\varepsilon + 1)$, there exist points $a' \in A \cap \mathbb{B}_\delta(a)$, $b' \in B \cap \mathbb{B}_\delta(b)$, $a^* \in N_A(a')$ and $b^* \in N_B(b')$ satisfying conditions (9).*

(ii) *Assertion (ii) in Theorem 3 holds true.*

Note that the extremality in part (i) of Definition 3 can be considered as a special case of the local extremality in part (ii) of that definition with $\rho = \infty$. On the other hand, as the next proposition shows, the local extremality can be considered as a special case of the extremality for a special pair of 'localized' sets.

Proposition 7 (Extremality vs local extremality) *Suppose X is a normed linear space, $A, B \subset X$, $a \in A$ and $b \in B$. If the pair $\{A, B\}$ is locally extremal relative to a and b with some $\rho > 0$, then, for any $\rho' \in (0, \rho)$, the pair $\{A \cap \mathbb{B}_{\rho'}(a), B \cap \mathbb{B}_{\rho'}(b)\}$ is extremal relative to a and b.*

Proof Let the pair $\{A, B\}$ be locally extremal relative to a and b with some $\rho > 0$, and numbers $\varepsilon > 0$ and $\rho' \in (0, \rho)$ be given. Choose an $\varepsilon' \in (0, \min\{\varepsilon, \rho - \rho'\})$. Then there exist $u, v \in X$ satisfying conditions (21) with ε' in place of ε. Hence, $\max\{\|u\|, \|v\|\} < \varepsilon' < \varepsilon$, $\rho'\mathbb{B} - u \subset \rho\mathbb{B}$, $\rho'\mathbb{B} - v \subset \rho\mathbb{B}$, and

$$(A \cap \mathbb{B}_{\rho'}(a) - a - u) \cap (B \cap \mathbb{B}_{\rho'}(b) - b - v)$$
$$= (A - a - u) \cap (\rho'\mathbb{B} - u) \cap (B - b - v) \cap (\rho'\mathbb{B} - v)$$
$$\subset (A - a - u) \cap (B - b - v) \cap (\rho\mathbb{B}) = \emptyset.$$

Thus, the pair $\{A \cap \mathbb{B}_{\rho'}(a), B \cap \mathbb{B}_{\rho'}(b)\}$ is extremal relative to a and b. $\qquad\square$

In view of Proposition 6, the next proposition shows that all the properties in Definition 3 are meaningful only when $a \in \mathrm{bd}\,A$ and $b \in \mathrm{bd}\,B$.

Proposition 8 (Approximate stationarity relative to boundary points) *Suppose X is a normed linear space, $A, B \subset X$, $a \in A$ and $b \in B$. If the pair $\{A, B\}$ is approximately stationary relative to a and b, then $a \in \mathrm{bd}\,A$ and $b \in \mathrm{bd}\,B$.*

Proof Suppose, on the contrary, that $a \in \mathrm{int}\,A$. (The case $b \in \mathrm{int}\,B$ is not much different.) Then $\mathbb{B}_r(a) \subset A$ for some $r > 0$. Choose an $\varepsilon \in (0, \min\{r/3, 1\})$. If $\rho \in (0, \varepsilon)$, $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ and $u, v \in (\varepsilon\rho)\mathbb{B}$, then $\hat{x} := -v \in (B - b' - v) \cap (\rho\mathbb{B})$ and $\|\hat{x} + a' + u - a\| \le \|a' - a\| + \|u\| + \|v\| < \varepsilon(1 + 2\rho) < 3\varepsilon < r$. Hence, $\hat{x} + a' + u \in A$ and $\hat{x} \in A - a' - u$. It follows that $(A - a' - u) \cap (B - b' - v) \cap (\rho\mathbb{B}) \ne \emptyset$, and consequently, the pair $\{A, B\}$ is not approximately stationary relative to a and b. $\qquad\square$

The conventional Definitions 1 and 2 of the extremality and stationarity properties of $\{A, B\}$ involve the translations $A - u$ and $B - v$ of the sets and, thus, refer implicitly to the 'relative' versions of the corresponding properties. The next proposition is in a sense a reformulation of Definition 3.

Proposition 9 (Conventional vs relative extremality and stationarity) *Suppose X is a normed linear space, $A, B \subset X$ and $\bar{x} \in A \cap B$. The pair $\{A, B\}$ is extremal/locally extremal/stationary/approximately stationary at $\bar{x}$ if and only if, for any $u, v \in X$, the pair $\{A - u, B - v\}$ is extremal/locally extremal/stationary/approximately stationary relative to $\bar{x} - u$ and $\bar{x} - v$.*

In the case of (local) extremality, the distance between the translated sets $A - u$ and $B - v$ does not have to be attained at the translated points $\bar{x} - u$ and $\bar{x} - v$; see Fig. 3.

Fig. 3 Conventional vs relative extremality

The conventional Definitions 1 and 2 of the extremality and stationarity properties as well as their extensions in Definition 3 involve vectors u, v determining the "shifts" of each of the sets. It was observed in [21] that in the case of conventional extremality and stationarity of two sets it is sufficient to shift one of the sets only. (In the general case of n ($n > 1$) sets, one can consider shifts of $n - 1$ sets.)

Proposition 10 (Relative extremality and stationarity with a single set shifted) *Suppose X is a normed linear space, $A, B \subset X$, $a \in A$ and $b \in B$. The pair $\{A, B\}$ is extremal/locally extremal/stationary/approximately stationary relative to a and b if and only if the respective conditions in Definition 3 are satisfied with $v = 0$.*

Proof If any of the conditions in Definition 3 is satisfied with $v = 0$, then the respective property obviously holds.

Conversely, if $(A - a - u) \cap (B - b - v) \cap (\rho \mathbb{B}) = \emptyset$ for some $\rho \in (0, \infty]$ and some $u, v \in X$, then $(A - a - u') \cap (B - b) \cap (\rho \mathbb{B} + v) = \emptyset$ where $u' := u - v$. Set $\rho' := \rho/2$. If $\max\{\|u\|, \|v\|\} < \alpha \leq \rho/2$, then $\|u'\| < 2\alpha$ and $\rho' \mathbb{B} \subset \rho \mathbb{B} + v$. Hence,

$$(A - a - u') \cap (B - b) \cap (\rho' \mathbb{B}) = \emptyset. \tag{22}$$

These simple observations allow one to ensure each of the conditions in Definition 3 with this u' and appropriate choice of ε and ρ.

Let condition (ii) in Definition 3 be satisfied with some $\rho \in (0, \infty]$. (As observed above, the case $\rho = \infty$ covers condition (i) in Definition 3.) Then, with $\rho' := \rho/2$ and any $\varepsilon' > 0$, one can take $\alpha := \varepsilon := \min\{\varepsilon', \rho\}/2$ and find $u, v \in X$ such that conditions (21) hold. With u' defined as above, one has $\|u'\| < 2\alpha \leq \varepsilon'$ and condition (22) is satisfied; hence, condition (ii) (condition (i) if $\rho = \infty$) in Definition 3 is satisfied with $v = 0$.

Let condition (iii) (condition (iv)) in Definition 3 be satisfied. Then, with any $\varepsilon' > 0$, one can take $\varepsilon := \min\{\varepsilon'/4, 1/2\}$ and find a $\rho \in (0, \varepsilon)$ and points $u, v \in X$ (and $a' \in A \cap \mathbb{B}_\varepsilon(a)$ and $b' \in B \cap \mathbb{B}_\varepsilon(b)$) such that conditions (14) hold. With $\rho' := \rho/2$, $\alpha := \varepsilon \rho \leq \rho/2$ and u' defined as above, one has $\rho' \in (0, \varepsilon')$, $\|u'\| < 2\alpha \leq \varepsilon' \rho'$ (and $a' \in A \cap \mathbb{B}_{\varepsilon'}(a)$ and $b' \in B \cap \mathbb{B}_{\varepsilon'}(b)$) and condition (22) is satisfied (with a' and b' in place of a and b); hence, condition (iii) (condition (iv)) in Definition 3 is satisfied with $v = 0$. $\qquad \square$

Remark 13 1. Condition $v = 0$ in Proposition 10 can be replaced with $u = 0$.

2. Similarly to Proposition 10, one can also impose condition $v = 0$ (or $u = 0$) in Theorem 6. However, unlike Proposition 10, this would require appropriate amendments in the estimates.

3. Thanks to Proposition 10, the relative extremality in Definition 3(i) is equivalent to the condition $0 \in \mathrm{bd}[(A - a) - (B - b)]$.

It was observed in [21, Theorem 1] (see also [22, Theorem 1]) that approximate stationarity in Definition 2(ii) can be characterized in metric terms. The next proposition provides a version of this result for the relative approximate stationarity in Definition 3(iv). It is a consequence of [22, Theorem 1(ii)] and Definition 3(iv).

Proposition 11 (Metric characterization of approximate stationarity) *Suppose X is a normed linear space, $A, B \subset X$, $a \in A$ and $b \in B$. The pair $\{A, B\}$ is approximately stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exist $y \in \mathbb{B}_\varepsilon(a)$, $z \in \mathbb{B}_\varepsilon(b)$ and $x \in \varepsilon\mathbb{B}$ such that*

$$\max\{d(x, A - y), d(x, B - z)\} < \varepsilon d(x, (A - y) \cap (B - z)).$$

The next proposition shows that, if the distance between A and B is attained (at least locally) at some points a and b, then the pair $\{A, B\}$ is (locally) extremal relative to a and b.

Proposition 12 (Relative extremality when the distance is attained) *Suppose X is a normed linear space, $A, B \subset X$ are closed, $a \in A$ and $b \in B$.*

(i) *If $\|a - b\| = d(A, B) > 0$, then the pair $\{A, B\}$ is extremal relative to a and b.*
(ii) *If $\|a - b\| = d(A \cap \mathbb{B}_\rho(a), B \cap \mathbb{B}_\rho(b)) > 0$ for some $\rho > 0$, then the pair $\{A, B\}$ is locally extremal relative to a and b.*

Proof (i) Take an arbitrary $\varepsilon > 0$ and a $t \in (0, \min\{\varepsilon/\|a - b\|, 1/2\})$. Set $u := t(b - a)$ and $v := t(a - b)$. Then $\max\{\|u\|, \|v\|\} = t\|a - b\| < \varepsilon$, and we only need to show that

$$(A - a - u) \cap (B - b - v) = \emptyset. \tag{23}$$

Suppose this is not true, i.e, there exists an $x \in X$ such that $a' := a + u + x \in A$ and $b' := b + v + x \in B$. Then $\|a' - b'\| = (1 - 2t)\|a - b\| < \|a - b\|$, which contradicts the assumption. Hence, condition (23) is true and, consequently, the pair $\{A, B\}$ is extremal relative to a and b.

(ii) Take a $\rho' \in (0, \rho)$, an arbitrary $\varepsilon > 0$ and a $t \in (0, \min\{\varepsilon/\|a - b\|, (\rho - \rho')/\|a - b\|, 1/2\})$. Set $u := t(b - a)$ and $v := t(a - b)$. Then $\max\{\|u\|, \|v\|\} = t\|a - b\| < \varepsilon$, and we only need to show that

$$(A - a - u) \cap (B - b - v) \cap (\rho'\mathbb{B}) = \emptyset. \tag{24}$$

Suppose this is not true, i.e, there exists an $x \in \rho'\mathbb{B}$ such that $a' := a + u + x \in A$ and $b' := b + v + x \in B$. Observe that $\max\{\|a' - a\|, \|b' - b\|\} \le t\|a - b\| + \rho' < \rho$. Thus, $a' \in A \cap \mathbb{B}_\rho(a)$ and $b' \in B \cap \mathbb{B}_\rho(b)$, and $\|a' - b'\| = (1 - 2t)\|a - b\| < \|a - b\|$, which contradicts the assumption. Hence, condition (23) is true and, consequently, the pair $\{A, B\}$ is locally extremal relative to a and b. $\qquad\square$

Remark 14 As the next example shows, the condition $d(A,B) > 0$ in part (i) and the similar condition in part (ii) of Proposition 12 cannot be dropped.

Example 5 For the sets $A := \{(x_1, x_2) \mid x_2 \leq 0\}$ and $B := \{(x_1, x_2) \mid x_1 + x_2 \leq 0\}$ in $\mathbb{R}^2$ we have $d(A,B) = 0$, while $\{A,B\}$ is obviously not extremal (in fact it is not even approximately stationary) at any $a = b \in A \cap B$ (in particular, at $a = b = 0$).

Remark 15 When $A \cap B = \emptyset$ and the distance is not attained, there may or may not be a pair of points $a \in A$ and $b \in B$ such that the pair $\{A,B\}$ is (locally) extremal relative to a and b.

Example 6 1. For the sets $A := \{(x_1, x_2) \mid x_2 \leq 0\}$ and $B := \{(x_1, x_2) \mid x_2 \geq e^{-x_1}\}$ in $\mathbb{R}^2$ (see Fig. 4) we have $A \cap B = \emptyset$, while $\{A,B\}$ is obviously not extremal (and even not approximately stationary) at any $a \in A$ and $b \in B$. Note that $\{A,B\}$ could still be considered extremal if the conventional Definition 1(i) was amended as discussed in Remark 5.

2. For the sets in item 1 above, it holds $d(A,B) = 0$. This is not a precondition. If the set B above is translated upwards by one unit: $B := \{(x_1, x_2) \mid x_2 \geq e^{-x_1} + 1\}$ (see Fig. 5), then $d(A,B) = 1$ and $\{A,B\}$ is still not extremal at any $a \in A$ and $b \in B$.

3. If the two sets in item 1 above are modified slightly: $A := \{(x_1, x_2) \mid -1 \leq x_2 \leq 0\}$ and $B := \{(x_1, x_2) \mid e^{-x_1} \leq x_2 \leq 1\}$ (see Fig. 6), then despite their intersection still being empty, the modified sets are extremal, e.g., at $a := (1, -1) \in A$ and $b := (1, 1) \in B$.
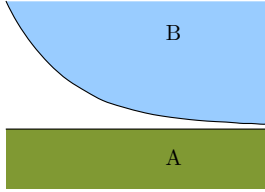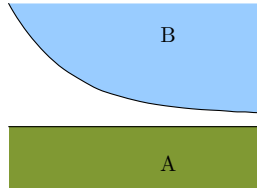
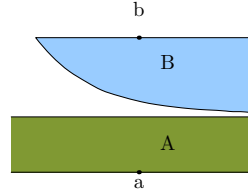

Fig. 4 Example 6.1   **Fig. 5** Example 6.2   **Fig. 6** Example 6.3

The approximate stationarity property in Definition 3(iii) possesses certain stability: if it holds at a certain pair (a,b), it holds 'approximately' at all nearby pairs.

Proposition 13 (Stability of relative approximate stationarity) *Suppose X is a normed linear space, $A,B \subset X$ are closed, $a \in A$ and $b \in B$. The pair $\{A,B\}$ is approximately stationary relative to a and b if and only if for any $\varepsilon > 0$, $\delta > 0$ and any points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ there exist $\rho \in (0, \delta)$, $a'' \in A \cap \mathbb{B}_\varepsilon(a')$, $b'' \in B \cap \mathbb{B}_\varepsilon(b')$ and $u, v \in X$ such that conditions (14) hold true with a'' and b'' in place of a and b.*

Proof The sufficiency is obvious: taking $a' := a$ and $b' := b$ in the conditions of Proposition 13, one satisfies the conditions in Definition 3(iii). To prove the necessity, let the pair $\{A,B\}$ be approximately stationary relative to a and b, $\varepsilon > 0$, $\delta > 0$, $a' \in A \cap \mathbb{B}_\varepsilon(a)$ and $b' \in B \cap \mathbb{B}_\varepsilon(b)$. Choose a $\xi \in (0, \delta)$ such that $\max\{\|a' - a\|, \|b' -$

$b\|\} + \xi < \varepsilon$. By Definition 3, there exist $\rho \in (0,\xi)$, $a'' \in A \cap \mathbb{B}_\xi(a)$, $b'' \in B \cap \mathbb{B}_\xi(b)$ and $u,v \in X$ such that conditions (14) hold true with a'' and b'' in place of a and b. Then $\rho < \delta$, $\|a'' - a'\| \leq \|a'' - a\| + \|a - a'\| < \varepsilon$ and, similarly, $\|b'' - b'\| < \varepsilon$. $\square$

4.3 Pairs of Sets and Set-valued Mappings

It is well known (see e.g. [10,29]) that regularity/transversality properties of collections of sets are in a sense equivalent to the corresponding regularity and Lipschitz-like properties of certain set-valued mappings. Given two subsets A and B of a normed linear space X, the mappings $F : X \rightrightarrows X \times X$ and $S : X \times X \rightrightarrows X$ defined by

$$F(x) := (A - x) \times (B - x), \ \ x \in X, \quad \text{and} \quad S(y,z) := (A - y) \cap (B - z), \ \ y,z \in X \quad (25)$$

play a major role in this type of analysis. Below we establish links between the relative extremality and stationarity properties of the pair of sets $\{A,B\}$ and certain properties of the set-valued mappings F and S given by (25).

First notice that $S = F^{-1}$, $\mathrm{dom}\, S = F(X)$, $F(0) = A \times B$ and $S(0,0) = A \cap B$. Given an $a \in A$ and a $b \in B$, we obviously have $(a,b) \in F(0)$ and $0 \in S(a,b)$. Recall that in this paper we are assuming that the product space $X \times X$ is equipped with the maximum norm. This corresponds to $\max\{\|u\|, \|v\|\}$ involved in all parts of Definition 3, and is not a big restriction: any other norm compatible with the norm on X can be used instead, as long as it is consistently used everywhere (cf. Remark 6).

Proposition 14 (Pairs of sets and set-valued mappings) *Suppose X is a normed linear space, $A,B \subset X$, $a \in A$, $b \in B$ and set-valued mappings F and S are given by (25).*

(i) *$\{A,B\}$ is extremal relative to a and b if and only if $(a,b) \in \mathrm{bd}\, F(X)$, or equivalently, $(a,b) \in \mathrm{bd}\,\mathrm{dom}\, S$.*
(ii) *$\{A,B\}$ is locally extremal relative to a and b if and only if there exists a $\rho > 0$ such that*
   (a) *$(a,b) \in \mathrm{bd}\, F(\rho\mathbb{B})$, or equivalently,*
   (b) *for any $\varepsilon > 0$, there is a pair $(y,z) \in \mathbb{B}_\varepsilon(a,b)$ such that $S(y,z) \cap (\rho\mathbb{B}) = \emptyset$.*
(iii) *$\{A,B\}$ is stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exists a $\rho \in (0,\varepsilon)$ such that*
   (a) *$\mathbb{B}_{\varepsilon\rho}(a,b) \not\subset F(\rho\mathbb{B})$, or equivalently,*
   (b) *there is a pair $(y,z) \in \mathbb{B}_{\varepsilon\rho}(a,b)$ such that $S(y,z) \cap (\rho\mathbb{B}) = \emptyset$.*
(iv) *$\{A,B\}$ is approximately stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exist a $\rho \in (0,\varepsilon)$ and points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ such that*
   (a) *$\mathbb{B}_{\varepsilon\rho}(a',b') \not\subset F(\rho\mathbb{B})$, or equivalently,*
   (b) *there is a pair $(y,z) \in \mathbb{B}_{\varepsilon\rho}(a',b')$ such that $S(y,z) \cap (\rho\mathbb{B}) = \emptyset$.*

Proof (i) By Definition 3(i), $\{A,B\}$ is extremal relative to a and b if and only if, for any $\varepsilon > 0$, there is a pair $(y,z) \in \mathbb{B}_\varepsilon(a,b)$ such that $S(y,z) = (A - y) \cap (B - z) = \emptyset$, i.e., $(a,b) \in \mathrm{bd}\,\mathrm{dom}\, S = \mathrm{bd}\, F(X)$.

(ii) By Definition 3(ii), $\{A,B\}$ is locally extremal relative to a and b if and only if there exists a $\rho > 0$ such that for any $\varepsilon > 0$ there is a pair $(y,z) \in \mathbb{B}_\varepsilon(a,b)$ such that $S(y,z) \cap (\rho\mathbb{B}) = (A-y) \cap (B-z) \cap (\rho\mathbb{B}) = \emptyset$. This is equivalent to $(a,b) \in \mathrm{bd}\, F(\rho\mathbb{B})$.

(iii) By Definition 3(iii), $\{A,B\}$ is stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exist a $\rho \in (0,\varepsilon)$ and a pair $(y,z) \in \mathbb{B}_{\varepsilon\rho}(a,b)$ such that $S(y,z) \cap (\rho\mathbb{B}) = (A-y) \cap (B-z) \cap (\rho\mathbb{B}) = \emptyset$. This is equivalent to $\mathbb{B}_{\varepsilon\rho}(a,b) \not\subset F(\rho\mathbb{B})$.

(iv) By Definition 3(iv), $\{A,B\}$ is approximately stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exist a $\rho \in (0,\varepsilon)$, points $a' \in A \cap \mathbb{B}_\varepsilon(a)$, $b' \in B \cap \mathbb{B}_\varepsilon(b)$ and a pair $(y,z) \in \mathbb{B}_{\varepsilon\rho}(a',b')$ such that $S(y,z) \cap (\rho\mathbb{B}) = (A-y) \cap (B-z) \cap (\rho\mathbb{B}) = \emptyset$. This is equivalent to $\mathbb{B}_{\varepsilon\rho}(a',b') \not\subset F(\rho\mathbb{B})$. □

Remark 16 The extremality and stationarity properties of pairs of sets studied in the current paper are in a sense examples of their 'irregular behaviour'; cf. [21,22,24]. No surprise, the properties of the set-valued mappings F and S that appear in Proposition 14 are in fact negations of certain regularity, semicontinuity and Lipschitz-like properties, some of which are well known. Below we briefly comment on these properties.

(i) In accordance with Proposition 14(i), $\{A,B\}$ *is NOT extremal relative to a and b if and only if there exists an $\alpha > 0$ such that* $\mathbb{B}_\alpha(a,b) \subset F(X)$*, or equivalently,* $\mathbb{B}_\alpha(a,b) \subset \mathrm{dom}\, S$. This means that F *covers* (a,b) (on X).

(ii) In accordance with Proposition 14(ii), $\{A,B\}$ *is NOT locally extremal relative to a and b if and only if, for any $\rho > 0$, there exists an $\alpha > 0$ such that*
(a) $\mathbb{B}_\alpha(a,b) \subset F(\rho\mathbb{B})$*, or equivalently,*
(b) $d(0,S(y,z)) \le \rho$ *for any* $(y,z) \in \mathbb{B}_\alpha(a,b)$.
This means that F *covers* (a,b) on $\rho\mathbb{B}$ (is *open* at $(0,(a,b))$ [5, p. 180]) and S is lower semicontinuous at $((a,b),0)$ [12, p. 10].

(iii) In accordance with Proposition 14(iii), $\{A,B\}$ *is NOT stationary relative to a and b if and only if there exists an $\alpha > 0$ such that*
(a) $\mathbb{B}_{\alpha\rho}(a,b) \subset F(\rho\mathbb{B})$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$, or equivalently,*
(b) $d(0,S(y,z)) \le \rho$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$,* $(y,z) \in \mathbb{B}_{\alpha\rho}(a,b)$.
Condition (a) means that F α-*covers* [23, p. 1765] (is α-*open* [1, Definition 2.4(i)]) at $(0,(a,b))$. It is equivalent [23, Theorem 6(i)] to the following condition:
(c) $\alpha d(0,F^{-1}(y,z)) \le d((y,z),(a,b))$ *for all (y,z) near (a,b),*
which means that F is *semiregular* [23, p. 1765], [4, Definition 1.2] (*hemiregular* [1, Definition 2.4(iii)]) at $(0,(a,b))$ with rank α.
In its turn, condition (b) can be rewritten as
(d) $\alpha d(0,S(y,z)) \le d((y,z),(a,b))$ *for all (y,z) near (a,b),*
which means that S is *Lipschitz lower semicontinuous* [12, p. 34] (*pseudocalm* [1, Definition 2.4(ii)]) at $(0,(a,b))$ with rank α.

(iv) In accordance with Proposition 14(iv), $\{A,B\}$ *is NOT approximately stationary relative to a and b if and only if there exists an $\alpha > 0$ such that*
(a) $\mathbb{B}_{\alpha\rho}(a',b') \subset F(\rho\mathbb{B})$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$, $a' \in A \cap \mathbb{B}_\delta(a)$, $b' \in B \cap \mathbb{B}_\delta(b)$, or equivalently,*
(b) $d(0,S(y,z)) \le \rho$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$, $a' \in A \cap \mathbb{B}_\delta(a)$, $b' \in B \cap \mathbb{B}_\delta(b)$, $(y,z) \in \mathbb{B}_{\alpha\rho}(a',b')$.*

It is not difficult to check that the above conditions are equivalent, respectively, to the following two (with possibly a smaller δ):

(a$'$) $\mathbb{B}_{\alpha\rho}(y,z) \subset F(\mathbb{B}_\rho(w))$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$, $(w,y,z) \in \mathrm{gph}\,F \cap$*
$\mathbb{B}_\delta(0,a,b)$*, or equivalently,*

(b$'$) $d(w,S(y',z')) \leq \rho$ *for some $\delta > 0$ and all $\rho \in (0,\delta)$, $(w,y,z) \in \mathrm{gph}\,F \cap$*
$\mathbb{B}_\delta(0,a,b)$*, $(y',z') \in \mathbb{B}_{\alpha\rho}(y,z)$.*

Condition (a$'$) means that F *uniformly covers* [23, p. 1766] (is *open with/at linear rate* [12, p. 13], [1, Definition 2.1(i)]) around $(0,(a,b))$. It is known to be equivalent (see e.g. [23, Theorem 6(iii)]) to the following condition:

(c) $\alpha d(w,F^{-1}(y,z)) \leq d((y,z),F(w))$ *for all (y,z) near (a,b) and w near 0,*

which means that F is *metrically regular* [12, p. 12], [5, p. 178] at $(0,(a,b))$ with rank α.

In its turn, condition (b) can be rewritten as

(d) $\alpha d(w,S(y',z')) \leq \|(y',z') - (y,z)\|$ *for all (y,z) and (y',z') near (a,b) and $w \in S(y,z)$ near 0,*

which means that S has the *Aubin property* [5, p. 172] (is *pseudo Lipschitz* [12, (D1)]) at $(0,(a,b))$ with rank α.

There is little consistency in the literature about whether to put α in the left or the right-hand side of the corresponding inequality/inclusion in the definitions of the properties discussed above. Thus, in some sources it is α^{-1} that is taken as the quantitative estimate (rank, modulus) of the respective property instead of α.

Observe from (25) that $F(x) = A \times B - (x,x)$. This simple observation provides a link between the extremality and stationarity properties of pairs of sets studied in the current paper and the *nonconvex separation property* introduced by Borwein and Jofre [2]. Given subsets $A, B \subset X$ and points $a \in A$ and $b \in B$, define another pair of sets and a pair of points

$$\tilde{A} := A \times B, \quad \tilde{B} := \{(x,x) \mid x \in X\} \quad \text{and} \quad \tilde{a} := (a,b) \in \tilde{A}, \quad \tilde{b} := (0,0) \in \tilde{B} \quad (26)$$

in the product space $X \times X$. Note that $\tilde{B} = -\tilde{B}$. As previously, this space is assumed to be equipped with the maximum norm. We are going to keep the standard notation $\mathbb{B}$ for the unit ball in $X \times X$. The next proposition is a consequence of Proposition 14.

Proposition 15 (Pairs of sets in X and $X \times X$) *Suppose X is a normed linear space, $A, B \subset X$, $a \in A$, $b \in B$ and the sets $\tilde{A}$ and $\tilde{B}$ and points $\tilde{a}$ and $\tilde{b}$ are given by* (26).

(i)   *$\{A,B\}$ is extremal relative to a and b if and only if $\tilde{a} + \tilde{b} \in \mathrm{bd}(\tilde{A} + \tilde{B})$.*

(ii) *$\{A,B\}$ is locally extremal relative to a and b if and only if there exists a $\rho > 0$ such that $\tilde{a} + \tilde{b} \in \mathrm{bd}(\tilde{A} + \tilde{B} \cap (\rho\mathbb{B}))$.*

(iii) *$\{A,B\}$ is stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exists a $\rho \in (0,\varepsilon)$ such that $\mathbb{B}_{\varepsilon\rho}(\tilde{a} + \tilde{b}) \not\subset \tilde{A} + \tilde{B} \cap (\rho\mathbb{B})$.*

(iv) *$\{A,B\}$ is approximately stationary relative to a and b if and only if, for any $\varepsilon > 0$, there exist a $\rho \in (0,\varepsilon)$ and a point $\tilde{a}' \in \tilde{A} \cap \mathbb{B}_\varepsilon(\tilde{a})$ such that $\mathbb{B}_{\varepsilon\rho}(\tilde{a}' + \tilde{b}) \not\subset \tilde{A} + \tilde{B} \cap (\rho\mathbb{B})$.*

The condition $\tilde{a} + \tilde{b} \in \mathrm{bd}(\tilde{A} + \tilde{B})$ in Proposition 26(i) is the *boundary condition* for the sets $\tilde{A}$ and $\tilde{B}$ introduced and characterized in [2, Theorem 1], while the assertions

in Proposition 26(i) and (ii) in the special case $a = b$ improve [2, Proposition 2(ii)] (see [21, Remark 9]). The conditions in parts (iii) and (iv) of Proposition 26 define certain stationarity properties for the sets $\tilde{A}$ and $\tilde{B}$ which may be of independent interest.

Remark 17 Proposition 26 relates the extremality and stationarity properties of pairs of sets in X with the corresponding 'boundary condition'-like properties of certain pairs of sets in $X \times X$. There is also a way in the opposite direction: given subsets $A, B \subset X$ and points $a \in A$ and $b \in B$, one can consider the boundary condition $a + b \in \mathrm{bd}(A + B)$ and its analogues as in the corresponding parts of Proposition 26 and relate them with the corresponding extremality and stationarity properties of the sets $\tilde{A} := A \times B$ and $\tilde{B} := \{(y, z) \mid y + z = a + b\}$ in $X \times X$ at the point (a, b) along the lines of [2, Proposition 2(i)] and [21, Section 4]).

References

1. Apetrii, M., Durea, M., Strugariu, R.: On subregularity properties of set-valued mappings. Set-Valued Var. Anal. **21**, 93–126 (2013)
2. Borwein, J.M., Jofré, A.: A nonconvex separation property in Banach spaces. Math. Methods Oper. Res. **48**, 169–179 (1998)
3. Borwein, J.M., Zhu, Q.J.: Techniques of Variational Analysis. Springer, New York (2005)
4. Cibulka, R., Fabian M., Kruger A.Y.: On semiregularity of mappings. arXiv:1711.04420 (2017)
5. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings: A View from Variational Analysis, 2nd edn. Springer Series in Operations Research and Financial Engineering. Springer, New York (2014)
6. Dubovitskii, A.Ya., Milyutin, A.A.: Extremum problems in the presence of restrictions. USSR Comp. Maths. Math. Phys. **5**, 1–80 (1965)
7. Fabian, M.: Subdifferentials, local ε-supports and Asplund spaces. J. Lond. Math. Soc. **s2-34**, 568–576 (1986)
8. Fabian, M.: Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. Acta Univ. Carol. **30**, 51–56 (1989)
9. Ioffe, A.D.: Fuzzy principles and characterization of trustworthiness. Set-Valued Anal. **6**, 265–276 (1998)
10. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russ. Math. Surveys **55**, 501–558 (2000)
11. Ioffe, A.D.: Transversality in variational analysis. J. Optim. Theory Appl. **174**, 343–366 (2017)
12. Klatte, D., Kummer, B.: Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications. Nonconvex Optimization and its Applications, vol. 60. Kluwer Academic Publishers, Dordrecht (2002)
13. Kruger, A.Y.: Generalized differentials of nonsmooth functions. VINITI no. 1332-81. Minsk (1981), 67 pp. (In Russian). Available at: https://asterius.ballarat.edu.au/akruger/research/publications.html
14. Kruger, A.Y.: ε-semidifferentials and ε-normal elements. VINITI no. 1331-81. Minsk (1981), 76 pp. (In Russian). Available at https://asterius.ballarat.edu.au/akruger/research/publications.html
15. Kruger, A.Y.: Generalized differentials of nonsmooth functions and necessary conditions for an extremum. Sibirsk. Mat. Zh. **26**, 78–90 (1985). In Russian; English transl.: Sib. Math. J. **26**, 370–379 (1985)
16. Kruger, A.Y.: About extremality of systems of sets. Dokl. Nats. Akad. Nauk Belarusi **42**, 24–28 (1998). In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html

17. Kruger, A.Y.: Strict (ε, δ)-semidifferentials and extremality of sets and functions. Dokl. Nats. Akad. Nauk Belarusi **44**, 19–22 (2000). In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html
18. Kruger, A.Y.: Strict (ε, δ)-subdifferentials and extremality conditions. Optimization **51**, 539–554 (2002)
19. Kruger, A.Y.: On Fréchet subdifferentials. J. Math. Sci. **116**, 3325–3358 (2003)
20. Kruger, A.Y.: Weak stationarity: eliminating the gap between necessary and sufficient conditions. Optimization **53**, 147–164 (2004)
21. Kruger, A.Y.: Stationarity and regularity of set systems. Pac. J. Optim. **1**, 101–126 (2005)
22. Kruger, A.Y.: About regularity of collections of sets. Set-Valued Anal. **14**, 187–206 (2006)
23. Kruger, A.Y.: About stationarity and regularity in variational analysis. Taiwan. J. Math. **13**, 1737–1785 (2009)
24. Kruger, A.Y., López, M.A.: Stationarity and regularity of infinite collections of sets. J. Optim. Theory Appl. **154**, 339–369 (2012)
25. Kruger, A.Y., Luke, D.R., Thao, N.H.: About subtransversality of collections of sets. Set-Valued Var. Anal. **25**, 701–729 (2017)
26. Kruger, A.Y., Mordukhovich, B.S.: Extremal points and the Euler equation in nonsmooth optimization problems. Dokl. Akad. Nauk BSSR **24**(8), 684–687 (1980). In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html
27. Kruger, A.Y., Mordukhovich, B.S.: Generalized normals and derivatives and necessary conditions for an extremum in problems of nondifferentiable programming. II. VINITI no. 494-80, 60 pp. Minsk (1980). In Russian. Available at https://asterius.ballarat.edu.au/akruger/research/publications.html
28. Kruger, A.Y., Thao, N.H.: About uniform regularity of collections of sets. Serdica Math. J. **39**, 287–312 (2013)
29. Kruger, A.Y., Thao, N.H.: Quantitative characterizations of regularity properties of collections of sets. J. Optim. Theory Appl. **164**, 41–67 (2015)
30. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation. I: Basic Theory. Grundlehren der Mathematischen Wissenschaften, vol. 330. Springer, Berlin Heidelberg (2006)
31. Mordukhovich, B.S., Shao, Y.: Extremal characterizations of Asplund spaces. Proc. Am. Math. Soc. **124**, 197–205 (1996)
32. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability. Lecture Notes in Mathematics, vol. 1364, 2nd edn. Springer-Verlag, Berlin Heidelberg (1993)
33. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin Heidelberg (1998)
34. Pierra, G.: Decomposition through formalization in a product space. Math. Program. **28**, 96–115 (1984)
35. Zheng, X.Y., Ng, K.F.: The Lagrange multiplier rule for multifunctions in Banach spaces. SIAM J. Optim. **17**, 1154–1175 (2006)
36. Zheng, X.Y., Ng, K.F.: A unified separation theorem for closed sets in a Banach space and optimality conditions for vector optimization. SIAM J. Optim. **21**, 886–911 (2011)
37. Zheng, X.Y., Yang, Z., Zou, J.: Exact separation theorem for closed sets in Asplund spaces. Optimization **66**, 1065–1077 (2017)