STATIONARITY AND REGULARITY OF REAL-VALUED FUNCTIONS

ALEXANDER YA. KRUGER

Centre of Information and Applied Optimization
School of Information Technology and Mathematical Sciences
University of Ballarat, Australia
a.kruger@ballarat.edu.au

Abstract

Different stationarity and regularity concepts for extended real-valued functions on metric spaces are considered in the paper. The properties are characterized in terms of certain local constants. A classification scheme for stationarity/regularity constants and corresponding concepts is proposed. The relations between different constants are established.

Key words: Variational analysis, subdifferential, normal cone, optimality, stationarity, regularity, slope, set-valued mapping, Asplund space.

1 Introduction

The papers considers different stationarity and regularity concepts for extended real-valued functions on metric spaces.

All the properties are characterized in terms of certain local constants. A function is said to be stationary at a point (in some sense) if the corresponding constant is zero (a critical point). Otherwise the function is said to be regular at this point (in the same sense) and the constant provides a quantitative estimate of regularity.

All the variety of constants and corresponding stationary/regularity concepts can be classified in the following way. Firstly, there are "inf" constants and concepts (characterizing a function from below and appropriate for minimization problems) and "sup" ones (characterizing a function from above and appropriate for maximization problems). One can also consider "combined" concepts. Combined stationary means that either an "inf" or a "sup" stationary condition is satisfied, while combined regularity corresponds to the case when both "inf" and "sup" regularity conditions hold true.

Secondly, there are "basic" constants (defined at a point) and "strict" or "fuzzy" ones (accumulating information about the function properties at nearby points). The latter constants lead to weak stationary and strong regularity concepts.

Thirdly, there are "primal" and "dual" constants (defined in terms of primal and dual space elements respectively) and corresponding stationary/regularity concepts. Dual constants can be defined when the primal space is a normed linear space.

For Fréchet differentiable or convex functions all stationary/regularity concepts reduce to traditional ones.

The definitions of the constants, the relations between them and the corresponding stationary and regularity concepts developed in the current paper are very similar to those for multifunctions and collections of sets (see [12, 13, 14]). Actually it is another application of the same variational approach which was used earlier for characterizing other types of objects. This confirms the assertion formulated by Jonathan M. Borwein and Qiji J. Zhu: "an important feature of the new variational techniques is that they can handle nonsmooth functions, sets and multifunctions equally well" [1].

The paper is organized as follows. The basic primal constants characterizing local properties of a function from below and corresponding inf-stationarity (inf-θ-stationarity) and inf-regularity (inf-θ-regularity) concepts are introduced and investigated in Section 2. Section 3 is devoted to the strict primal inf-type constants and weak inf-stationarity and strong inf-regularity concepts. It is proved that in the case of a lower semicontinuous function on a complete metric space two different basic constants produce the same strict constant. Sup-type as well as combined constants (both basic and strict) and corresponding stationarity and regularity concepts are considered in Section 4. Sections 5 and 6 are devoted to primal constants for Fréchet differentiable and convex functions respectively. Dual (subdifferential) stationarity and regularity conditions (in the case of a normed linear space) are formulated in Section 7 in terms of Fréchet subdifferentials and strict δ-subdifferentials. In the case of an Asplund space these conditions appear to be equivalent to the corresponding strict primal conditions.

Mainly standard for variational analysis notations (see the books [1, 16, 18]) are used throughout the paper. A closed ball of radios ρ centered at x in a metric space is denoted by $B_\rho(x)$. We write B_ρ if $x = 0$, and simply B if $x = 0$ and $\rho = 1$.

2 Inf-Stationarity and Inf-Regularity

Let φ be a function defined on a metric space X with values in an extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. It is assumed to be finite at some point $x^\circ \in X$.

For $\rho > 0$ define the constant

$$\theta_\rho[\varphi](x^\circ) = \inf_{x \in B_\rho(x^\circ)} \varphi(x) - \varphi(x^\circ). \tag{1}$$

This constant as well as the behavior of the function $\rho \to \theta_\rho[\varphi](x^\circ)$ near 0 can be used for characterizing local properties of φ near x°.

Proposition 1. *(i) $\theta_\rho[\varphi](x^\circ) \le 0$ for all $\rho > 0$;*

(ii) $\rho \to \theta_\rho[\varphi](x^\circ)$ is nonincreasing on $\mathbb{R}_+$.

Proposition 2. *φ attains at x° a*

(i) global minimum if and only if $\theta_\rho[\varphi](x^\circ) = 0$ for all $\rho > 0$;

(ii) local minimum if and only if $\theta_\rho[\varphi](x^\circ) = 0$ for some $\rho > 0$.

Proposition 3. *φ is lower semicontinuous at x° if and only if $\lim_{\rho \to +0} \theta_\rho[\varphi](x^\circ) = 0$.*

The next step is to "differentiate" the function $\rho \to \theta_\rho[\varphi](x^\circ)$ at 0 (from the right). Define

$$\theta[\varphi](x^\circ) = \limsup_{\rho \to +0} \frac{\theta_\rho[\varphi](x^\circ)}{\rho}. \tag{2}$$

This is also a nonpositive constant. The case $\theta[\varphi](x^\circ) = 0$ corresponds to a kind of stationary behavior of φ near x°.

Definition 1. φ is

(i) *inf-θ-stationary* at x° if $\theta[\varphi](x^\circ) = 0$;

(ii) *inf-θ-regular* at x° if $\theta[\varphi](x^\circ) < 0$.

Remark 1. The purpose of the "inf" prefix in Definition 1 is to emphasize that minimization problems are addressed here[1]. Contrary to the classical case stationarity-regularity properties of nondifferentiable functions "from below" and "from above" can be essentially different.

Another derivative-like constant that can be used for characterizing stationarity-regularity properties of φ at x° can be defined in the following way:

$$\tau[\varphi](x^\circ) = \liminf_{x \to x^\circ} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{d(x, x^\circ)}. \tag{3}$$

The notation $[\alpha]_- = \min(\alpha, 0)$ is used here. Similar to (2) the constant (3) is nonnegative.

Definition 2. φ is

(i) *inf-stationary* at x° if $\tau[\varphi](x^\circ) = 0$;

(ii) *inf-regular* at x° if $\tau[\varphi](x^\circ) < 0$.

The relations between (2) and (3) are given by the next proposition.

Proposition 4. *The following assertions hold true:*

(i) $\tau[\varphi](x^\circ) \le \theta[\varphi](x^\circ)$*;*

(ii) if $\theta_\rho[\varphi](x^\circ) = 0$ *for some* $\rho > 0$ *then* $\tau[\varphi](x^\circ) = \theta[\varphi](x^\circ) = 0$.

Proof. (i) Let $\rho > 0$ be arbitrary. Then

$$\frac{\theta_\rho[\varphi](x^\circ)}{\rho} = \inf_{x \in B_\rho(x^\circ)} \frac{\varphi(x) - \varphi(x^\circ)}{\rho} = \inf_{x \in B_\rho(x^\circ)} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\rho} = \inf_{x \in B_\rho(x^\circ) \setminus \{x^\circ\}} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\rho} \ge \inf_{x \in B_\rho(x^\circ) \setminus \{x^\circ\}} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{d(x, x^\circ)}.$$

The assertion follows now from (2) and (3).

(ii) If $\theta_\rho[\varphi](x^\circ) = 0$ for some $\rho > 0$ then by definition (1) $\varphi(x) \ge \varphi(x^\circ)$ for all $x \in B_\rho(x^\circ)$ and it follows from (3) that $\tau[\varphi](x^\circ) = 0$. □

Thus, using (3) instead of (2) leads to a stronger concept of stationarity and correspondingly to a weaker concept of regularity.

Corollary 4.1. *The following assertions hold true:*

(i) If φ *is inf-stationary at* x° *then it is inf-*θ*-stationary at* x°*;*

(ii) If φ *is inf-*θ*-regular at* x° *then it is inf-regular at* x°*;*

(iii) If φ *attains a local minimum at* x° *then it is both inf-stationary and inf-*θ*-stationary at* x°*.*

Inequality (i) in Proposition 4 can be strict even for functions from $\mathbb{R}$ to $\mathbb{R}$.

Example 1. Take $\varphi(x) = -|x|$, if $|x| = 1/2^n$, $n = 1, 2, \ldots$, and $\varphi(x) = 0$ otherwise. Obviously $\tau[\varphi](0) = -1$. At the same time, for any $\rho \in \Xi_n = \{\rho : 1/2^n \le \rho < 1/2^{n-1}\}$ one has $\theta_\rho[\varphi](0) = -1/2^n$ and

$$\sup_{\rho \in \Xi_n} \frac{\theta_\rho[\varphi](0)}{\rho} = \frac{-1/2^n}{1/2^{n-1}} = -\frac{1}{2}.$$

Thus, $\theta[\varphi](0) = -1/2$.

[1]The terminology was suggested by V. F. Demianov (personal communication).

It is possible to modify the above example to make $\theta[\varphi](0)$ equal zero.

Example 2. Take $\varphi(x) = -|x|$, if $|x| = 1/n^n$, $n = 1, 2, \ldots$, and $\varphi(x) = 0$ otherwise. One still has $\tau[\varphi](0) = -1$ while $\theta[\varphi](0) = 0$.

Thus, in the above example φ is inf-regular at 0 while being inf-θ-stationary at this point.

It is possible to modify the example further to make φ continuous and even differentiable near 0 (but not strictly differentiable!) while keeping the inequality (i) in Proposition 4 strict.

Remark 2. $\tau[\varphi](x^\circ)$ coincides up to a sign with the *slope* $|\nabla\varphi|(x^\circ)$ of φ at x° [3] (see also [7]).

3 Weak Inf-Stationarity and Strong Inf-Regularity

Definitions (2) and (3) can be modified further by allowing x° in their right-hand sides to vary near a given point. In that way we arrive at two more nonpositive derivative-like constants:

$$\hat{\theta}[\varphi](x^\circ) = \limsup_{\substack{x \xrightarrow{\varphi} x^\circ \\ \rho \to +0}} \inf_{u \in B_\rho(x)} \frac{\varphi(u) - \varphi(x)}{\rho}, \tag{4}$$

$$\hat{\tau}[\varphi](x^\circ) = \limsup_{\substack{x \xrightarrow{\varphi} x^\circ \\ \rho \to +0}} \inf_{u \in B_\rho(x) \setminus \{x\}} \frac{[\varphi(u) - \varphi(x)]_-}{d(u, x)}. \tag{5}$$

Notation $x \xrightarrow{\varphi} x^\circ$ here means that $x \to x^\circ$ with $\varphi(x) \to \varphi(x^\circ)$. Due to variations of x (4) and (5) gain some properties of a strict derivative. They are used below for defining some more stationarity and regularity concepts.

Definition 3. φ is

(i) *weakly inf-θ-stationary* at x° if $\hat{\theta}[\varphi](x^\circ) = 0$;

(ii) *strongly inf-θ-regular* at x° if $\hat{\theta}[\varphi](x^\circ) < 0$;

(iii) *weakly inf-stationary* at x° if $\hat{\tau}[\varphi](x^\circ) = 0$;

(iv) *strongly inf-regular* at x° if $\hat{\tau}[\varphi](x^\circ) < 0$.

The next proposition summarizes some interrelations between the constants above.

Theorem 1. *The following assertions hold true:*

(i) $\hat{\theta}[\varphi](x^\circ) \geq \limsup_{x \xrightarrow{\varphi} x^\circ} \theta[\varphi](x)$,

(ii) $\hat{\tau}[\varphi](x^\circ) = \limsup_{x \xrightarrow{\varphi} x^\circ} \tau[\varphi](x)$;

(iii) $\hat{\tau}[\varphi](x^\circ) \leq \hat{\theta}[\varphi](x^\circ)$;

(iv) If X is complete and φ is lower semicontinuous near x°, then $\hat{\tau}[\varphi](x^\circ) = \hat{\theta}[\varphi](x^\circ)$.

Proof. (i) By the definition of the lower limit one can write

$$\hat{\theta}[\varphi](x^\circ) = \lim_{\delta \to +0} \sup_{\substack{x \in B_\delta(x^\circ) \\ |\varphi(x) - \varphi(x)| \leq \delta \\ 0 < \rho \leq \delta}} \frac{\theta_\rho[\varphi](x)}{\rho} = \lim_{\delta \to +0} \sup_{\substack{x \in B_\delta(x^\circ) \\ |\varphi(x) - \varphi(x)| \leq \delta}} \sup_{0 < \rho \leq \delta} \frac{\theta_\rho[\varphi](x)}{\rho}. \tag{6}$$

The legality of the replacement of the "double" supremum in the above formula by two separate ones is quite obvious. For any $x \in X$ and any $0 < \delta' \le \delta$ one has

$$\sup_{0<\rho\le\delta} \frac{\theta_\rho[\varphi](x)}{\rho} \ge \sup_{0<\rho\le\delta'} \frac{\theta_\rho[\varphi](x)}{\rho}.$$

Consequently,

$$\sup_{0<\rho\le\delta} \frac{\theta_\rho[\varphi](x)}{\rho} \ge \lim_{\delta'\to+0} \sup_{0<\rho\le\delta'} \frac{\theta_\rho[\varphi](x)}{\rho} = \limsup_{\rho\to+0} \frac{\theta_\rho[\varphi](x)}{\rho} = \theta[\varphi](x). \tag{7}$$

The assertion follows from (6) and (7).

(ii) In a similar way

$$\hat{\tau}[\varphi](x^\circ) = \lim_{\delta\to+0} \sup_{\substack{x\in B_\delta(x^\circ) \\ |\varphi(x)-\varphi(x)|\le\delta}} \sup_{0<\rho\le\delta} \inf_{u\in B_\rho(x)\setminus\{x\}} \frac{[\varphi(u)-\varphi(x)]_-}{d(u,x)} =$$

$$\limsup_{x\xrightarrow{\varphi}x^\circ} \lim_{\rho\to+0} \inf_{u\in B_\rho(x)\setminus\{x\}} \frac{[\varphi(u)-\varphi(x)]_-}{d(u,x)} = \limsup_{x\xrightarrow{\varphi}x^\circ} \tau[\varphi](x).$$

(iii) follows from (i) and (ii) due to part (i) of Proposition 4.

(iv) Let X be complete and φ be lower semicontinuous near x°. Due to (iii) we need to show the opposite inequality[2]. Denote $\alpha = -\hat{\theta}[\varphi](x^\circ) \ge 0$ and take arbitrary $\varepsilon > 0$. Then it follows from (4) that there exists an $x' \in B_{\varepsilon/2}(x^\circ)$ with $|\varphi(x') - \varphi(x^\circ)| \le \varepsilon/2$ and a $\rho > 0$, satisfying $\rho \le (\varepsilon/2)\min(1, (\alpha+\varepsilon/2)^{-1})$, such that

$$\varphi(u) - \varphi(x') \ge -(\alpha+\varepsilon/2)\rho$$

for any $u \in B_\rho(x')$. Without loss of generality we can assume that φ is lower semicontinuous on $B_\rho(x')$ and apply *Ekeland variational principle* [4]. Take some ρ' satisfying

$$\rho\frac{\alpha+\varepsilon/2}{\alpha+\varepsilon} \le \rho' < \rho.$$

Then there exists an $x \in B_{\rho'}(x')$ such that $\varphi(x) \le \varphi(x')$ and

$$\varphi(u) - \varphi(x) + (\alpha+\varepsilon/2)(\rho/\rho')d(u,x) \ge 0$$

for all u near x. Thus,

$$\tau[\varphi](x) \ge -(\alpha+\varepsilon/2)(\rho/\rho') \ge -(\alpha+\varepsilon),$$

while $x \in B_{\rho'}(x') \subset B_\varepsilon(x^\circ)$ since $d(x', x^\circ) \le \varepsilon/2$ and $\rho' < \rho \le \varepsilon/2$. At the same time

$$\varphi(x) - \varphi(x^\circ) \le \varphi(x') - \varphi(x^\circ) \le \varepsilon/2,$$
$$\varphi(x) - \varphi(x^\circ) = \varphi(x) - \varphi(x') + \varphi(x') - \varphi(x^\circ) \ge -(\alpha+\varepsilon/2)\rho - \varepsilon/2 \ge -\varepsilon.$$

Consequently, $|\varphi(x) - \varphi(x^\circ)| \le \varepsilon$. Since ε is arbitrary, taking into account (ii), one can conclude that $\hat{\tau}[\varphi](x^\circ) \ge -\alpha = \hat{\theta}[\varphi](x^\circ)$. □

Parts (i) and (ii) of Theorem 1 imply the inequalities

$$\theta[\varphi](x^\circ) \le \hat{\theta}[\varphi](x^\circ), \quad \tau[\varphi](x^\circ) \le \hat{\tau}[\varphi](x^\circ),$$

and they can be strict.

[2] The proof follows the original idea suggested by Bernd Kummer for the case $\hat{\theta}[\varphi](x^\circ) = 0$ (personal communication).

Example 3. Take the function φ from Example 1. Evidently, φ attains a local minimum at $x_n = 1/2^n$ for any $n = 1, 2, \ldots$. Due to Proposition 2 $\theta_\rho[\varphi](x_n) = 0$ for some $\rho > 0$, and it follows from Proposition 4 that $\tau[\varphi](x_n) = \theta[\varphi](x_n) = 0$. Consequently, $\hat{\tau}[\varphi](0) = \hat{\theta}[\varphi](0) = 0$. Recall that $\tau[\varphi](0) = -1$ and $\theta[\varphi](0) = -1/2$.

Inequalities (i) and (iii) in Theorem 1 can be strict too.

Example 4. Define the function $\varphi : \mathbb{R} \to \mathbb{R}$ in the following way: $\varphi(x) = x$ if $x \leq 0$, $\varphi(x) = x - 1/n$ if $1/n < x \leq 1/(n-1)$, $n = 2, 3, \ldots$, $\varphi(x) = x - 1/2$ if $x > 1/2$. It is easy to see that $\theta[\varphi](x) = \tau[\varphi](x) = -1$ for any $x \in \mathbb{R}$. Then $\hat{\tau}[\varphi](0) = -1$. On the other hand, take $x_n = 1/n + 1/n^2$, $\rho_n = 1/n$, $n = 1, 2, \ldots$. Then $\varphi(x_n) = 1/n^2$ and consequently $\theta_{\rho_n}[\varphi](x_n) \geq -1/n^2$. It follows immediately that $\hat{\theta}[\varphi](0) = 0$.

Remark 3. Note that due to part (iv) of Theorem 1 in the case of a lower semicontinuous function on a complete metric space two different constants (2) and (3) produce (in accordance with (4) and (5)) the same "strict" constant.

Corollary 1.1. *The following assertions hold true:*

(i) If φ is inf-θ-stationary (inf-stationary) at x° then it is weakly inf-θ-stationary (weakly inf-stationary) at x°;

(ii) If φ is strongly inf-θ-regular (strongly inf-regular) at x° then it is inf-θ-regular (inf-regular) at x°;

(iii) If φ is weakly inf-stationary at x° then it is weakly inf-θ-stationary at x°;

(iv) If φ is strongly inf-θ-regular at x° then it is strongly inf-regular at x°.

Let, additionally, X be complete and φ be lower semicontinuous near x°. Then

(v) φ is weakly inf-stationary at x° if and only if it is weakly inf-θ-stationary at x°;

(vi) φ is strongly inf-θ-regular at x° if and only if it is strongly inf-regular at x°.

The next "fuzzy" characterization of weak inf-stationarity can be convenient for applications. It follows directly from definition (5).

Proposition 5. *φ is weakly inf-stationary at x° if and only if for any $\varepsilon > 0$ there exists an $x \in B_\varepsilon(x^\circ)$ such that $|\varphi(x) - \varphi(x^\circ)| \leq \varepsilon$ and $\varphi(u) + \varepsilon d(u, x) \geq \varphi(x)$ for all u near x.*

Remark 4. A point x satisfying $\varphi(u) + \varepsilon d(u, x) \geq \varphi(x)$ for all u near x is referred to in [15] (see also [8]) as a *local Ekeland point* of φ (with factor ε). If all the conditions in Proposition 5 are satisfied then x° is said to be a *stationary point* of φ *with respect to minimization* [15]. Thus, stationarity with respect to minimization is equivalent to weak inf-stationarity and, in case of a lower semicontinuous function on a complete metric space, also to weak inf-θ-stationarity.

4 Other Types of Stationarity and Regularity

Similarly to (1)–(5) corresponding "maximization" constants can be defined. To do this one has to replace "inf", "lim inf" and $[\cdot]_-$ in (1), (3) by "sup", "lim sup" and $[\cdot]_+$ respectively, and "lim sup" in (2), (4), (5) by "lim inf". The resulting constants appear to be nonnegative. They are related to (1)–(5) by the following equalities:

$$\theta^+_\rho[\varphi](x^\circ) = -\theta_\rho[-\varphi](x^\circ), \quad \theta^+[\varphi](x^\circ) = -\theta[-\varphi](x^\circ), \quad \tau^+[\varphi](x^\circ) = -\tau[-\varphi](x^\circ),$$
$$\hat{\theta}^+[\varphi](x^\circ) = -\hat{\theta}[-\varphi](x^\circ), \quad \hat{\tau}^+[\varphi](x^\circ) = -\hat{\tau}[-\varphi](x^\circ)$$

and lead to similar *sup-stationarity* and *sup-regularity* concepts.

Of course, for a function φ the set of sup-stationary (sup-regular) points is different in general from that of inf-stationary (inf-regular) points.

Example 5. Take the function φ from Example 1. One has $\tau[\varphi](0) < \theta[\varphi](0) < 0$. Thus, φ is both inf-regular and inf-θ-regular at 0. At the same time 0 is a point of maximum and $\tau^+[\varphi](0) = \theta^+[\varphi](0) = 0$. Consequently, φ is both sup-stationarity and sup-θ-stationarity at 0.

The "combined" concepts can also be of interest. It is natural to say that a function is stationary (in some sense) at a point if it is either inf-stationary or sup-stationary at this point. On the contrary, the regularity property for a function is satisfied when this function is both inf-regular and sup-regular at the point.

Definition 4. φ is

(i) *θ-stationary* at x° if $\max(\theta[\varphi](x^\circ), \theta[-\varphi](x^\circ)) = 0$;

(ii) *θ-regular* at x° if $\max(\theta[\varphi](x^\circ), \theta[-\varphi](x^\circ)) < 0$.

(iii) *stationary* at x° if $\max(\tau[\varphi](x^\circ), \tau[-\varphi](x^\circ)) = 0$;

(iv) *regular* at x° if $\max(\tau[\varphi](x^\circ), \tau[-\varphi](x^\circ)) < 0$;

(v) *weakly θ-stationary* at x° if $\max(\hat{\theta}[\varphi](x^\circ), \hat{\theta}[-\varphi](x^\circ)) = 0$;

(vi) *strongly θ-regular* at x° if $\max(\hat{\theta}[\varphi](x^\circ), \hat{\theta}[-\varphi](x^\circ)) < 0$;

(vii) *weakly stationary* at x° if $\max(\hat{\tau}[\varphi](x^\circ), \hat{\tau}[-\varphi](x^\circ)) = 0$;

(viii) *strongly regular* at x° if $\max(\hat{\tau}[\varphi](x^\circ), \hat{\tau}[-\varphi](x^\circ)) < 0$.

The relations between the concepts from Definition 4 are very similar to those formulated in Sections 2 and 3. In particular, strong regularity and strong θ-regularity coincide if X is complete and φ is continuous near x°.

Strong inf-regularity can be interpreted in the following way: all points in a neighborhood of a given point have "descent sequences", and the rate of descent is uniform. In contrast to that, strong regularity is equivalent to the existence of both descent and ascent sequences with the uniformity property.

Proposition 6. *(i) φ is strongly inf-regular at x° if and only if there exists an $\alpha > 0$ and a $\delta > 0$ such that for any $x \in B_\delta(x^\circ)$ with $|\varphi(x) - \varphi(x^\circ)| \leq \delta$ and any $\rho \in (0, \delta]$ one can find $u \in B_\rho(x)$ such that $\varphi(u) - \varphi(x) < -\alpha d(u, x)$.*

(ii) φ is strongly regular at x° if and only if there exists an $\alpha > 0$ and a $\delta > 0$ such that for any $x \in B_\delta(x^\circ)$ with $|\varphi(x) - \varphi(x^\circ)| \leq \delta$ and any $\rho \in (0, \delta]$ one can find $u', u'' \in B_\rho(x)$ such that $\varphi(u') - \varphi(x) < -\alpha d(u', x)$ and $\varphi(u'') - \varphi(x) > \alpha d(u'', x)$.

5 Differentiable Functions

The constants and corresponding stationarity/regularity concepts defined in the preceding sections take quite a traditional form when the function is assumed differentiable or convex (see Section 6). Fortunately, the number of different constants and concepts reduces significantly.

We will assume in the rest of the paper that X is a normed linear space with the distance induced by the norm.

Theorem 2. *If φ is Fréchet differentiable at x° with the derivative $\nabla\varphi(x^\circ)$ then*

$$\theta[\varphi](x^\circ) = \tau[\varphi](x^\circ) = -\theta^+[\varphi](x^\circ) = -\tau^+[\varphi](x^\circ) = -\|\nabla\varphi(x^\circ)\|.$$

If, additionally, the derivative is strict then

$$\hat{\theta}[\varphi](x^\circ) = \hat{\tau}[\varphi](x^\circ) = -\hat{\theta}^+[\varphi](x^\circ) = -\hat{\tau}^+[\varphi](x^\circ) = -\|\nabla\varphi(x^\circ)\|.$$

Recall that φ is called *strictly differentiable* [2, 18] at x° (with the derivative $\nabla\varphi(x^\circ)$) if

$$\lim_{x\to x^\circ,\ u\to x^\circ} \frac{\varphi(u)-\varphi(x)-\langle\nabla\varphi(x^\circ),u-x\rangle}{\|u-x\|}=0.$$

Clearly this condition is stronger than traditional Fréchet differentiability. Thus, condition $\nabla\varphi(x^\circ)\neq 0$ does not guarantee strong regularity in the sense of Definition 3 (or Definition 4) unless φ is strictly differentiable at x°.

Example 6. Take $\varphi(x) = x + x^2\sin(1/x)$, if $x \neq 0$, and $\varphi(0) = 0$. This function is everywhere Fréchet differentiable and $\nabla\varphi(0) = 1$. Thus, φ is regular at zero. At the same time $\hat\tau[\varphi](0) = \hat\tau^+[\varphi](0) = 0$: there exists a sequence $x_k \to 0$ such that $\nabla\varphi(x_k) \to 0$, and the assertion follows from Theorem1, part (ii). Consequently, φ is both weakly inf-stationary and weakly sup-stationary at zero.

Proof of Theorem 2. Take arbitrary $\varepsilon > 0$. It follows from the definition of the Fréchet derivative that there exists a $\delta > 0$ such that

$$|\varphi(x)-\varphi(x^\circ)-\langle\nabla\varphi(x^\circ),x-x^\circ\rangle| \le \varepsilon\,\|x-x^\circ\| \tag{8}$$

for any $x \in B_\delta(x^\circ)$. If $0 < \rho \le \delta$ then

$$\left|\inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{\varphi(x)-\varphi(x^\circ)}{\|x-x^\circ\|} - \inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{\langle\nabla\varphi(x^\circ),x-x^\circ\rangle}{\|x-x^\circ\|}\right| \le \varepsilon.$$

It is easy to see that

$$\inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{\langle\nabla\varphi(x^\circ),x-x^\circ\rangle}{\|x-x^\circ\|} = -\,\|\nabla\varphi(x^\circ)\|$$

and

$$\inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{\varphi(x)-\varphi(x^\circ)}{\|x-x^\circ\|} = \inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{[\varphi(x)-\varphi(x^\circ)]_-}{\|x-x^\circ\|} \to \tau[\varphi](x^\circ)$$

as $\rho \to +0$. The equality in the last expression follows from the next simple observation: if $\varphi(x) > \varphi(x^\circ)$ for all $x \in B_\rho(x^\circ)\setminus\{x^\circ\}$ (this is the only case when the $[\cdot]_-$ operation could make a difference) then x° is a point of local minimum, $\nabla\varphi(x^\circ) = 0$ and consequently

$$\inf_{x\in B_\rho(x^\circ)\setminus\{x^\circ\}} \frac{\varphi(x)-\varphi(x^\circ)}{\|x-x^\circ\|} = 0.$$

Thus, $|\tau[\varphi](x^\circ) + \|\nabla\varphi(x^\circ)\|| \le \varepsilon$. Taking into account that ε is an arbitrary positive number one can conclude that

$$\tau[\varphi](x^\circ) = -\,\|\nabla\varphi(x^\circ)\|$$

and

$$\tau^+[\varphi](x^\circ) = -\tau[-\varphi](x^\circ) = \|\nabla\varphi(x^\circ)\|\,.$$

Similarly, it follows from (8) that

$$|\varphi(x)-\varphi(x^\circ)-\langle\nabla\varphi(x^\circ),x-x^\circ\rangle| \le \varepsilon\rho$$

for any $x \in B_\rho(x^\circ)$ and consequently

$$|\theta_\rho[\varphi](x^\circ) + \|\nabla\varphi(x^\circ)\|\,\rho| = \left|\inf_{x\in B_\rho(x^\circ)}(\varphi(x)-\varphi(x^\circ)) - \inf_{x\in B_\rho(x^\circ)}\langle\nabla\varphi(x^\circ),x-x^\circ\rangle\right| \le \varepsilon\rho$$

and

$$|\theta[\varphi](x^\circ) + \|\nabla\varphi(x^\circ)\|| \le \varepsilon.$$

This implies

$$\theta[\varphi](x^\circ) = -\left\|\nabla\varphi(x^\circ)\right\|$$

and

$$\theta^+[\varphi](x^\circ) = -\theta[-\varphi](x^\circ) = \left\|\nabla\varphi(x^\circ)\right\|.$$

Let φ be strictly differentiable at x°. In this case instead of (8) one has a stronger condition: there exists a $\delta > 0$ such that

$$|\varphi(u) - \varphi(x) - \langle\nabla\varphi(x^\circ), u - x\rangle| \leq \varepsilon \left\|u - x\right\| \tag{9}$$

for any $x, u \in B_\delta(x^\circ)$. Take arbitrary $x \in B_{\delta/2}(x^\circ)$ and $\rho \in (0, \delta/2]$. Then (9) holds true for any $u \in B_\rho(x)$. As above, this leads to the estimates

$$\left|\inf_{u\in B_\rho(x)\setminus\{x\}} \frac{[\varphi(u) - \varphi(x)]_-}{d(u,x)} + \left\|\nabla\varphi(x^\circ)\right\|\right| \leq \varepsilon,$$
$$\left|\inf_{u\in B_\rho(x)} \frac{\varphi(u) - \varphi(x)}{\rho} + \left\|\nabla\varphi(x^\circ)\right\|\right| \leq \varepsilon$$

valid for any $x \in B_{\delta/2}(x^\circ)$. Consequently,

$$|\hat{\tau}[\varphi](x^\circ) + \left\|\nabla\varphi(x^\circ)\right\|| \leq \varepsilon, \quad \left|\hat{\theta}[\varphi](x^\circ) + \left\|\nabla\varphi(x^\circ)\right\|\right| \leq \varepsilon$$

and

$$\hat{\theta}[\varphi](x^\circ) = \hat{\tau}[\varphi](x^\circ) = -\left\|\nabla\varphi(x^\circ)\right\|.$$

The equalities

$$\hat{\theta}^+[\varphi](x^\circ) = \hat{\tau}^+[\varphi](x^\circ) = \left\|\nabla\varphi(x^\circ)\right\|$$

follow immediately. □

Corollary 2.1. *If φ is Fréchet differentiable at x° with the derivative $\nabla\varphi(x^\circ)$ then the following conditions are equivalent.*

(i) φ is inf-θ-stationary at x°;

(ii) φ is inf-stationary at x°;

(iii) φ is θ-stationary at x°;

(iv) φ is stationary at x°;

(v) $\nabla\varphi(x^\circ) = 0$.

If, additionally, the derivative is strict then the above conditions are also equivalent to the following ones.

(vi) φ is weakly inf-θ-stationary at x°;

(vii) φ is weakly inf-stationary at x°;

(viii) φ is weakly θ-stationary at x°;

(ix) φ is weakly stationary at x°.

Remark 5. Stationary and weak stationary in the above corollary can be replaced with regularity and strong regularity respectively if one replaces the equality in (v) with the inequality $\nabla\varphi(x^\circ) \neq 0$.

6 Convex Functions

In the convex case, as one could expect, all versions of inf-stationarity coincide and appear to be equivalent to just (local and global) minimality.

Theorem 3. *Let φ be convex.*

(i) If $\theta_\rho[\varphi](x^\circ) < 0$ for some $\rho > 0$ then $\theta_\rho[\varphi](x^\circ) < 0$ for all $\rho > 0$.

(ii) The functions $\rho \to \theta_\rho[\varphi](x^\circ)/\rho$, $\rho \to \theta^+_\rho[\varphi](x^\circ)/\rho$ are nondecreasing on $\mathbb{R}_+\backslash\{0\}$.

(iii) The following equalities hold true:

$$\hat{\theta}[\varphi](x^\circ) = \hat{\tau}[\varphi](x^\circ) = \theta[\varphi](x^\circ) = \tau[\varphi](x^\circ) = \inf_{\rho>0} \frac{\theta_\rho[\varphi](x^\circ)}{\rho} = \inf_{x\neq x^\circ} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\|x - x^\circ\|},$$

$$\theta^+[\varphi](x^\circ) = \tau^+[\varphi](x^\circ) = \inf_{\rho>0} \frac{\theta^+_\rho[\varphi](x^\circ)}{\rho} = \inf_{\rho>0} \sup_{\|x-x^\circ\|=\rho} \frac{[\varphi(x) - \varphi(x^\circ)]_+}{\rho}.$$

(iv) $\tau[\varphi](x^\circ) + \tau^+[\varphi](x^\circ) \geq 0$.

(v) $\hat{\tau}[\varphi](x^\circ) + \hat{\tau}^+[\varphi](x^\circ) \geq 0$.

(vi) If $\tau[\varphi](x^\circ) + \tau^+[\varphi](x^\circ) = 0$ and $\{x_k\} \subset X$ is a sequence defining $\tau[\varphi](x^\circ)$, i.e. $x_k \to 0$ and

$$\tau[\varphi](x^\circ) = \lim_{k\to\infty} \frac{\varphi(x^\circ + x_k) - \varphi(x^\circ)}{\|x_k\|} \tag{10}$$

then $\{-x_k\}$ is a sequence defining $\tau^+[\varphi](x^\circ)$:

$$\tau^+[\varphi](x^\circ) = \lim_{k\to\infty} \frac{\varphi(x^\circ - x_k) - \varphi(x^\circ)}{\|x_k\|}. \tag{11}$$

Proof. We need the following elementary property of a convex function:

$$\varphi(x') - \varphi(x^\circ) \leq t(\varphi(x) - \varphi(x^\circ)), \tag{12}$$

where $x, x' \in X$, $0 \leq t \leq 1$ and $x' = x^\circ + t(x - x^\circ)$.

Let $0 < \rho' \leq \rho$, $t = \rho'/\rho$. Evidently $x \in B_\rho(x^\circ)$ if and only if $x' \in B_{\rho'}(x^\circ)$. It follows from (12) and (1) that $\theta_{\rho'}[\varphi](x^\circ) \leq t\theta_\rho[\varphi](x^\circ)$ and $\theta^+_{\rho'}[\varphi](x^\circ) \leq t\theta^+_\rho[\varphi](x^\circ)$. In other words,

$$\theta_{\rho'}[\varphi](x^\circ)/\rho' \leq \theta_\rho[\varphi](x^\circ)/\rho, \quad \theta^+_{\rho'}[\varphi](x^\circ)/\rho' \leq \theta^+_\rho[\varphi](x^\circ)/\rho,$$

which proves (ii). If $\theta_\rho[\varphi](x^\circ) < 0$ for some $\rho > 0$ then the first of the above inequalities guarantees that $\theta_{\rho'}[\varphi](x^\circ) < 0$ for all positive $\rho' \leq \rho$. Certainly in this case one also has $\theta_{\rho'}[\varphi](x^\circ) < 0$ for all $\rho' > \rho$ since the function $\rho \to \theta_\rho[\varphi](x^\circ)$ is nonincreasing. This proves (i).

(iii) Due to (ii), the representations of $\theta[\varphi](x^\circ)$ and $\theta^+[\varphi](x^\circ)$ can be simplified:

$$\theta[\varphi](x^\circ) = \inf_{\rho>0} \frac{\theta_\rho[\varphi](x^\circ)}{\rho}, \quad \theta^+[\varphi](x^\circ) = \inf_{\rho>0} \frac{\theta^+_\rho[\varphi](x^\circ)}{\rho}. \tag{13}$$

Similarly, it follows from (12) that

$$\frac{\varphi(x') - \varphi(x^\circ)}{\|x' - x^\circ\|} \leq \frac{\varphi(x) - \varphi(x^\circ)}{\|x - x^\circ\|}$$

if $x \neq x^\circ$ and $x' \neq x^\circ$, which yields the representations

$$\tau[\varphi](x^\circ) = \inf_{x \neq x^\circ} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\|x - x^\circ\|}, \tag{14}$$

$$\tau^+[\varphi](x^\circ) = \inf_{\rho > 0} \sup_{\|x - x^\circ\| = \rho} \frac{[\varphi(x) - \varphi(x^\circ)]_+}{\rho}. \tag{15}$$

Let us show that $\theta[\varphi](x^\circ) = \tau[\varphi](x^\circ)$. Due to inequality (i) in Proposition 4 we only need to prove that $\theta[\varphi](x^\circ) \leq \tau[\varphi](x^\circ)$. From (13) and (1) one has

$$\theta[\varphi](x^\circ) \leq \frac{\theta_\rho[\varphi](x^\circ)}{\rho} \leq \frac{\varphi(x) - \varphi(x^\circ)}{\rho}$$

for any $\rho > 0$ and any $x \in B_\rho(x^\circ)$. In particular, for any $x \neq x^\circ$ and $\rho = \|x - x^\circ\|$ one gets the inequality

$$\theta[\varphi](x^\circ) \leq \frac{\varphi(x) - \varphi(x^\circ)}{\|x - x^\circ\|}$$

and consequently

$$\theta[\varphi](x^\circ) \leq \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\|x - x^\circ\|}$$

since $\theta[\varphi](x^\circ) \leq 0$. The desired inequality follows from (14).

The proof of the equality $\theta^+[\varphi](x^\circ) = \tau^+[\varphi](x^\circ)$ is straightforward. Proposition 4, (i) implies the inequality $\theta^+[\varphi](x^\circ) \leq \tau^+[\varphi](x^\circ)$. The opposite inequality $\tau^+[\varphi](x^\circ) \leq \theta^+[\varphi](x^\circ)$ follows from (15) and the definition of $\theta^+[\varphi](x^\circ)$.

$\theta[\varphi](x^\circ) = \tau[\varphi](x^\circ)$ and $\theta^+[\varphi](x^\circ) = \tau^+[\varphi](x^\circ)$ imply the equalities $\hat\theta[\varphi](x^\circ) = \hat\tau[\varphi](x^\circ)$ and $\hat\theta^+[\varphi](x^\circ) = \hat\tau^+[\varphi](x^\circ)$ by definitions (4) and(5).

Due to parts (ii) and (iii) of Theorem 1 one has

$$\theta[\varphi](x^\circ) = \tau[\varphi](x^\circ) \leq \hat\tau[\varphi](x^\circ) \leq \hat\theta[\varphi](x^\circ).$$

To complete the proof of (iii) we only need to show that $\hat\theta[\varphi](x^\circ) \leq \theta[\varphi](x^\circ)$. If $\theta[\varphi](x^\circ) = 0$ the assertion is trivial. Let $\theta[\varphi](x^\circ) < \alpha < 0$. It is sufficient to show that $\hat\theta[\varphi](x^\circ) \leq \alpha$. Take arbitrary $\beta \in (\theta[\varphi](x^\circ), \alpha)$. Then it follows from (2) that there exists a $\delta > 0$ such that $\theta_\delta[\varphi](x^\circ) < \beta\delta$, and consequently $\varphi(u) - \varphi(x^\circ) < \beta\delta$ for some $u \in B_\delta(x^\circ)$. Denote $\delta_1 = (\alpha - \beta)\delta$ and take arbitrary $x \in B_{\delta_1}(x^\circ)$ such that $|\varphi(x) - \varphi(x^\circ)| \leq \delta_1$. Then $\varphi(u) - \varphi(x) < \alpha\delta$, $u \in B_{\delta_2}(x)$, where $\delta_2 = (1 + \alpha - \beta)\delta$, and consequently

$$\inf_{u \in B_{\delta_2}(x)} \frac{\varphi(u) - \varphi(x)}{\delta_2} < \frac{\alpha}{1 + \alpha - \beta}.$$

Due to (ii) the last inequality implies that

$$\inf_{u \in B_\rho(x)} \frac{\varphi(u) - \varphi(x)}{\rho} < \frac{\alpha}{1 + \alpha - \beta}$$

for all positive $\rho \leq \delta_2$ and all $x \in B_{\delta_1}(x^\circ)$ such that $|\varphi(x) - \varphi(x^\circ)| \leq \delta_1$. Consequently, $\hat\theta[\varphi](x^\circ) \leq \alpha/(1 + \alpha - \beta)$. Since β can be taken arbitrarily close to α one can conclude that $\hat\theta[\varphi](x^\circ) \leq \alpha$.

(iv) If $\tau[\varphi](x^\circ) = 0$ the inequality holds trivially. Assume that $\tau[\varphi](x^\circ) < 0$ and chose a (defining) sequence $x_k \to 0$ such that (10) holds true. Then

$$(\varphi(x^\circ + x_k) + \varphi(x^\circ - x_k))/2 \geq \varphi(x^\circ) \tag{16}$$

due to convexity of φ. Evidently,

$$\tau^+[\varphi](x^\circ) \geq \limsup_{k\to\infty} \frac{\varphi(x^\circ - x_k) - \varphi(x^\circ)}{\|x_k\|}. \tag{17}$$

On the other hand, taking into account (16), one has

$$\liminf_{k\to\infty} \frac{\varphi(x^\circ - x_k) - \varphi(x^\circ)}{\|x_k\|} \geq -\lim_{k\to\infty} \frac{\varphi(x^\circ + x_k) - \varphi(x^\circ)}{\|x_k\|} = -\tau[\varphi](x^\circ). \tag{18}$$

(17) and (18) imply the desired inequality $\tau[\varphi](x^\circ) + \tau^+[\varphi](x^\circ) \geq 0$.

(v) follows from (iv) due to Theorem 1, part (ii):

$$\hat{\tau}^+[\varphi](x^\circ) = \liminf_{x \xrightarrow{\varphi} x^\circ} \tau^+[\varphi](x) \geq -\limsup_{x \xrightarrow{\varphi} x^\circ} \tau[\varphi](x) = -\hat{\tau}[\varphi](x^\circ).$$

(vi) Due to the assumption $\tau[\varphi](x^\circ) + \tau^+[\varphi](x^\circ) = 0$ the constants $\tau[\varphi](x^\circ)$ and $\tau^+[\varphi](x^\circ)$ are either both nonzero or both zero. In the latter case one obviously has

$$\lim_{x\to 0} \frac{\varphi(x^\circ + x) - \varphi(x^\circ)}{\|x\|} = 0$$

and consequently any sequence $x_k \to 0$ is a defining sequence (for any of the two constants). This explains why the $[\cdot]_-$ and $[\cdot]_+$ operations are omitted in (10) and (11). If $\{x_k\}$ is a sequence defining $\tau[\varphi](x^\circ)$ then due to (17) and (18) one has the estimates

$$\tau^+[\varphi](x^\circ) \geq \limsup_{k\to\infty} \frac{\varphi(x^\circ - x_k) - \varphi(x^\circ)}{\|x_k\|} \geq$$
$$\liminf_{k\to\infty} \frac{\varphi(x^\circ - x_k) - \varphi(x^\circ)}{\|x_k\|} \geq -\tau[\varphi](x^\circ) = \tau^+[\varphi](x^\circ).$$

which imply (11). □

Corollary 3.1. *If φ is convex then the following conditions are equivalent.*

(i) φ attains a global minimum at x°;

(ii) φ attains a local minimum at x°;

(iii) φ is inf-θ-stationary at x°;

(iv) φ is inf-stationary at x°;

(v) φ is θ-stationary at x°;

(vi) φ is stationary at x°;

(vii) φ is weakly inf-θ-stationary at x°;

(viii) φ is weakly inf-stationary at x°;

(ix) φ is weakly stationary at x°.

Proof. (i) ⇔ (ii) follows from Theorem 3, (i).

(i) ⇒ (iii) follows from the equality

$$\theta[\varphi](x^\circ) = \inf_{\rho>0} \frac{\theta_\rho[\varphi](x^\circ)}{\rho}$$

in Theorem 3, (iii). This equality implies also the estimate $\theta[\varphi](x^\circ) \leq \theta_\rho[\varphi](x^\circ)/\rho \leq 0$ for any $\rho > 0$ which in its turn yields the opposite implication (iii) ⇒ (i).

(iii) ⇔ (iv) ⇔ (vii) ⇔ (viii) follow from Theorem 3, (iii).

(iii) ⇒ (v), (iv) ⇒ (vi), (viii) ⇒ (ix) follow from Definition 4, while the opposite implications follow from Theorem 3, (iv) and (iii). □

Remark 6. Conditions $\tau[\varphi](x^\circ) = \tau^+[\varphi](x^\circ) = 0$ imply the Fréchet differentiability of φ at x° (with the derivative equal to zero). The weaker condition $\tau[\varphi](x^\circ) + \tau^+[\varphi](x^\circ) = 0$ in Theorem 3, (v) implies linearity of the directional derivative of φ along the direction of steepest descent (if the latter exists) with the opposite direction being automatically the direction of steepest ascent. This condition is not sufficient for differentiability of φ at x° unless $X = \mathbb{R}$. Note also that the direction opposite to the direction of steepest ascent does not need to be a direction of steepest descent.

Example 7 ([3]). Take the function $\varphi(x, y) = \max(x, y)$ on $\mathbb{R}^2$ and assume that $\mathbb{R}^2$ is equipped with the max type norm: $\|x, y\| = \max(|x|, |y|)$. φ is obviously not differentiable at 0. At the same time $\tau[\varphi](0) = -1$, $\tau^+[\varphi](0) = 1$. The vector $(-1, -1)$ defines the (unique) direction of steepest descent. The opposite vector $(1, 1)$ defines the direction of steepest ascent and φ is linear along the line defined by these vectors. Note that the direction of steepest ascent is not unique. For instance, the vector $(1, 0)$ also defines the direction of steepest ascent, while the opposite vector does not define the direction of steepest descent and φ is not linear along this line.

Remark 7. Stationary and weak stationary in the assertions (iii) – (ix) of the above corollary can be replaced with regularity and strong regularity respectively if one replaces (i) and (ii) with the opposite assertions: x° is not a point of (local or global) minimum of φ.

7 Subdifferential Conditions

The stationarity and regularity properties were defined above in terms of primal space elements. In the case of a normed linear space they admit some dual characterizations in terms of *Fréchet subdifferentials.*

Let X be a normed linear space. Its (topological) dual is denoted X^*. $\langle \cdot, \cdot \rangle$ is the bilinear form defining the duality pairing. Recall that the Fréchet subdifferential of φ at x° is defined as

$$\partial\varphi(x^\circ) = \left\{ x^* \in X^* : \liminf_{x \to x^\circ} \frac{\varphi(x) - \varphi(x^\circ) - \langle x^*, x - x^\circ \rangle}{\|x - x^\circ\|} \geq 0 \right\}. \tag{19}$$

The simplest dual characterization of stationarity is given by the next proposition.

Proposition 7. *φ is inf-stationary at x° if and only if* $0 \in \partial\varphi(x^\circ)$.

Proof. By definition (19) the condition $0 \in \partial\varphi(x^\circ)$ is equivalent to

$$\liminf_{x \to x^\circ} \frac{\varphi(x) - \varphi(x^\circ)}{\|x - x^\circ\|} \geq 0,$$

which in its turn can be rewritten as

$$\liminf_{x \to x^\circ} \frac{[\varphi(x) - \varphi(x^\circ)]_-}{\|x - x^\circ\|} = 0.$$

By (3) this is equivalent to $\tau[\varphi](x^\circ) = 0$. □

Remark 8. Due to Corollary 4.1 the inclusion $0 \in \partial\varphi(x^\circ)$ is sufficient for inf-θ-stationarity of φ at x°. The opposite implication is not true in general (see Examples 1 and 2).

In what follows φ is assumed to be lower semicontinuous near x°.

In general nonconvex setting the subdifferential mapping $\partial\varphi(\cdot)$ fails to posses good (semi-)continuity properties. In fact, the set $\partial\varphi(x)$ can be empty rather often. Based on (19) one can define a more robust derivative-like object:

[3]The example was composed by Alexander Rubinov (personal communication).

$$\hat{\partial}_\delta \varphi(x^\circ) = \bigcup \{\partial\varphi(x) :\ x \in B_\delta(x^\circ),\ |\varphi(x) - \varphi(x^\circ)| \le \delta\}\,. \tag{20}$$

It depends on a positive parameter δ and accumulates information on "differential" properties of φ at nearby points, thus attaining some properties of a strict derivative. The set (20) is called a *strict δ-subdifferential* of φ at x° (see [9, 10, 11]). In contrast to (19), the set (20) can be nonconvex. However, it possesses certain subdifferential calculus.

Using (20) one more constant can be defined for characterizing stationarity/regularity properties of φ:

$$\eta[\varphi](x^\circ) = \lim_{\delta\to 0} \inf \left\{ \|x^*\| :\ x^* \in \hat{\partial}_\delta \varphi(x^\circ) \right\}. \tag{21}$$

In contrast to the constants considered in the preceding sections this constant is nonnegative.

Definition 5. φ is

(i) *inf-η-stationary* at x° if $\eta[\varphi](x^\circ) = 0$;

(ii) *inf-η-regular* at x° if $\eta[\varphi](x^\circ) > 0$.

Remark 9. Note that the inf-η-stationary condition $\eta[\varphi](x^\circ) = 0$ does not imply the inclusion $0 \in \hat{\partial}_\delta \varphi(x^\circ)$.

Example 8. Take $\varphi(x) = x$, if $x < 0$, and $\varphi(x) = x^2$ otherwise. One has $\partial\varphi(0) = \emptyset$, $0 \notin \hat{\partial}_\delta \varphi(0)$ for any $\delta > 0$ while $\eta[\varphi](0) = 0$.

Fortunately (21) happens to be closely related to (4) (and (5)).

Theorem 4. *(i) $\hat{\theta}[\varphi](x^\circ) + \eta[\varphi](x^\circ) \ge 0$.*

(ii) If X is Asplund and $\hat{\theta}[\varphi](x^\circ) > -1$ then $\eta[\varphi](x^\circ) \le -\hat{\theta}[\varphi](x^\circ)/(1 + \hat{\theta}[\varphi](x^\circ))$.

This theorem follows from [12], Theorem 2. The first part of the theorem is elementary. The proof of the second part is based on the application of the two fundamental results of variational analysis: the *Ekeland variational principle* [4] and the *fuzzy sum rule* due to M. Fabian [5].

Recall that a Banach space is called *Asplund* (see [6, 17]) if any continuous convex function on it is Fréchet differentiable on a dense G_δ subset.

Corollary 4.1. *(i) If φ is inf-η-stationary at x° then it is weakly inf-stationary at x°.*

(ii) If φ is strongly inf-regular at x° then it is inf-η-regular at x°.

(iii) If X is Asplund then the conditions are equivalent: φ is weakly inf-stationary (strongly inf-regular) at x° if and only if it is inf-η-stationary (inf-η-regular) at x°.

Remark 10. The dual "maximization" constant can be defined in a natural way if one replaces the strict δ-subdifferential in (21) with the strict δ-superdifferential. It is related to (21) by the formula $\eta^+[\varphi](x^\circ) = \eta[-\varphi](x^\circ)$ and leads to corresponding sup-η-stationarity and sup-η-regularity concepts. Theorem 4 implies similar relations between $\eta^+[\varphi](x^\circ)$ and $\hat{\theta}^+[\varphi](x^\circ)$.

Acknowledgements. I wish to thank Alexander Rubinov for numerous helpful discussions.

References

[1] Borwein, J. M., and Zhu, Q. J. *Techniques of Variational Analysis.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 20. Springer-Verlag, New York, 2005.

[2] Bourbaki, N. *Variétés Différentielles et Analytiques.* Herman, Paris, 1967.

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

[4] Ekeland, I. On the variational principle. *J. Math. Anal. Appl. 47* (1974), 324–353.

[5] Fabian, M. Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. *Acta Univ. Carolinae 30* (1989), 51–56.

[6] Fabian, M. *Gâteaux Differentiability of Convex Functions and Topology. Weak Asplund Spaces.* Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1997.

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

[8] Klatte, D., and Kummer, B. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, vol. 60 of *Nonconvex Optimization and Its Applications.* Kluwer Academic Publishers, Dordrecht, 2002.

[9] Kruger, A. Y. On calculus of strict ε-semidifferentials. *Dokl. Akad. Nauk Belarusi 40*, 4 (1996), 34–39. in Russian.

[10] Kruger, A. Y. Strict (ε, δ)-semidifferentials and extremality conditions. *Optimization 51* (2002), 539–554.

[11] Kruger, A. Y. On Fréchet subdifferentials. *J. Math. Sci. (N. Y.) 116*, 3 (2003), 3325–3358. Optimization and related topics, 3.

[12] Kruger, A. Y. Weak stationarity: eliminating the gap between necessary and sufficient conditions. *Optimization 53* (2004), 147–164.

[13] Kruger, A. Y. Stationarity and regularity of set systems. *Pac. J. Optim. 1*, 1 (2005), 101–126.

[14] Kruger, A. Y. About regularity of collections of sets. *Set-Valued Anal.*, (2006), to appear.

[15] Kummer, B. Inverse functions of pseudo regular mappings and regularity conditions. *Math. Program., Ser. B 88* (2000), 313–339.

[16] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, vol. 330, 331 of *Grundlehren Series (Fundamental Principles of Mathematical Sciences).* Springer-Verlag, New York, 2005.

[17] Phelps, R. R. *Convex Functions, Monotone Operators and Differentiability, 2nd edition.* Lecture Notes in Mathematics, Vol. 1364. Springer-Verlag, New York, 1993.

[18] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis.* Springer-Verlag, New York, 1998.