

I. Ginchev

Weakened subdifferentials and Fréchet
differentiability of real functions

2008/3



UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

<http://eco.uninsubria.it>

In questi quaderni vengono pubblicati i lavori dei docenti della Facoltà di Economia dell'Università dell'Insubria. La pubblicazione di contributi di altri studiosi, che abbiano un rapporto didattico o scientifico stabile con la Facoltà, può essere proposta da un professore della Facoltà, dopo che il contributo sia stato discusso pubblicamente. Il nome del proponente è riportato in nota all'articolo. I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

These Working papers collect the work of the Faculty of Economics of the University of Insubria. The publication of work by other Authors can be proposed by a member of the Faculty, provided that the paper has been presented in public. The name of the proposer is reported in a footnote. The views expressed in the Working papers reflect the opinions of the Authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

© Copyright I. Ginchev
Printed in Italy in April 2008
Università degli Studi dell'Insubria
Via Monte Generoso, 71, 21100 Varese, Italy

All rights reserved. No part of this paper may be reproduced in any form without permission of the Author.

Weakened subdifferentials and Fréchet differentiability of real functions*

Ivan Ginchev[†]

Abstract

Let X be a real Banach space and $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$. It is well known that the Clarke subdifferential $\partial^\circ f(x)$ of the function f at $x \in \text{int dom } f$ is a singleton if and only if f is strongly differentiable (then $\partial^\circ f(x) = \{D_s f(x)\}$, where $D_s f(x)$ is the strong subdifferential of f at x). Simple examples show that there exist Fréchet differentiable at x functions f , for which $\partial^\circ f(x)$ is not a singleton. In such a sense the Clarke subdifferential is not an exact generalization of the differential of a differentiable function. In the present paper we propose a new subdifferential $\partial^w f(x)$, called the weakened subdifferential of f at x , which preserves the nice calculus rules of the Clarke subdifferential, and for X finite dimensional, is a singleton $\partial^w f(x) = \{\zeta\}$ if and only if f is Fréchet differentiable at x , and then $\zeta = D_F f(x)$.

Key words: generalized subdifferentials.

2000 Mathematics Subject Classification: 49J52.

1 Introduction

In this paper X denotes a real Banach space with norm $\|\cdot\|$, and X^* is its dual. The open unit ball in X is denoted by B . The set of the real numbers is denoted by \mathbb{R} . We put also $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

Let $f : X \rightarrow \overline{\mathbb{R}}$ be locally Lipschitz near $x \in \text{int dom } f$ and $v \in X$. The Clarke directional derivative of f at x in direction v is defined by

$$f^\circ(x, v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{1}{t} \Delta f(y, v, t), \quad (1)$$

where $\Delta f(y, v, t) = (f(y + tv) - f(y))$. The Clarke subdifferential $\partial^\circ f(x)$ of f at x is defined by

$$\partial^\circ f(x) = \{\xi \in X^* \mid \langle \xi, v \rangle \leq f^\circ(x, v) \text{ for all } v \in X\}.$$

*A shorten version is published as: I. Ginchev, Weakened subdifferentials and Fréchet differentiability of real functions, In: Mathematics and Informatics - Reality and Perspectives, 19–30, VSU “Chernorizets Hrabâr” Universitetsko Izdatelstvo, Varna 2007 (ISBN 978-954-715-376-9).

[†]University of Insubria, Department of Economics, Via Monte Generoso 71, 21100 Italy
e-mail: iginchev@eco.uninsubria.it.

The Clarke subdifferential is closely linked to the strict differentiability. We remind that the function $f : X \rightarrow \overline{\mathbb{R}}$ is strictly differentiable at $x \in \text{int dom } f$ if there exists an element $D_s f(x) \in X^*$, called the strict differential of f at x , such that $f^\circ(x, v) = \lim_{y \rightarrow x, t \downarrow 0} \frac{1}{t} \Delta f(y, v, t) = \langle D_s f(x), v \rangle$. The following theorem holds.

Theorem 1 (Clarke [2]). *Let $f : X \rightarrow \overline{\mathbb{R}}$, $x \in \text{int dom } f$ and $\zeta \in X^*$. Then the following assertions are equivalent:*

- a) f is strictly differentiable at x and $D_s f(x) = \zeta$,
- b) f is Lipschitz near x and $\partial^\circ f(x)$ is the singleton $\{\zeta\}$.

The following simple example shows that there exist Frèchet differentiable at x functions f , for which $\partial^\circ f(x)$ is not a singleton.

Example 1. *The function $f : X \rightarrow \mathbb{R}$ defined by*

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is Frèchet differentiable, but not strictly differentiable, at $x = 0$. The Frèchet differential is $Df(0) = 0$ and does not coincide with the Clarke subdifferential $\partial^\circ f(0) = [-1, 1]$.

The purpose of the present paper is to propose a new subdifferential $\partial^w f(x)$, called the weakened subdifferential of f at x , which preserves the good calculus rules of the Clarke subdifferential $\partial^\circ f(x)$ [2], and such that $\partial^w f(x)$ is the singleton $\{D_F f(x)\}$ if and only if there exists the Fréchet differential $D_F f(x)$. This task is fulfilled, at least in the finite-dimensional case (though the weakened subdifferential is defined for functions in arbitrary normed spaces). In the last section it is underlined that the inclusion $\partial^w f(x) \subset \partial^\circ f(x)$ makes the weakened subdifferential more sensitive than the Clarke subdifferential in applications to optimization problems. At the same time, in opposite to many existing in the literature generalized subdifferentials, see Aubin, Frankovska [1], Pshenichnyi [4], and elsewhere, it preserves the nice calculus rules of the Clarke subdifferential. This observation gives some advantage of the weakened subdifferentials with respect to other generalized subdifferentials, and motivates eventual further their investigation.

2 The weakened derivative

In this section we define the notion of the weakened directional derivative $f^w(x, v)$ of a given function $f : X \rightarrow \overline{\mathbb{R}}$ at a point $x \in \text{int dom } X$ in direction $v \in X$. The function f is not supposed to be Lipschitz near x . If $v \in X$ we put

$$f^w(x, v) = \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, v, t). \quad (2)$$

In this definition $f^w(x, v) \in \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

The following propositions give the basic properties of the weakened derivative.

Proposition 1. Let $f : X \rightarrow \overline{\mathbb{R}}$, $x \in \text{int dom } f$.

a) The function $f^w(x, \cdot) : X \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ is sublinear; i. e. positively homogeneous and subadditive.

b) If f is Lipschitz with constant K near x , so is $f^w(x, \cdot)$.

c) The following inequality is true

$$f^w(x, v) \leq f^\circ(x, v). \quad (3)$$

d) It holds

$$f^w(x, -v) = (-f)^w(x, v) = -\lim_{k \rightarrow \infty} \liminf_{t \downarrow 0} \inf_{y \in x + ktB} \frac{1}{t} \Delta f(y, v, t).$$

Proof. a) Let $\lambda > 0$ and suppose that $k > 0$ and $t > 0$ are fixed numbers. Put $\tau = \lambda t$, $\kappa = k/\lambda$. Then

$$\frac{1}{t} \Delta f(y, \lambda v, t) = \lambda \frac{1}{\lambda t} (f(y + t\lambda v) - f(y)) = \lambda \frac{1}{\tau} \Delta f(y, v, \tau).$$

We have also $x + ktB = x + (k/\lambda)\tau B$. Since $t \rightarrow 0^+$ is equivalent to $\tau \rightarrow 0^+$, and $k \rightarrow \infty$ to $\kappa \rightarrow \infty$, we get easily

$$\begin{aligned} f^w(x, \lambda v) &= \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, \lambda v, t) \\ &= \lambda \lim_{\kappa \rightarrow \infty} \limsup_{\tau \downarrow 0} \sup_{y \in x + \kappa \tau B} \frac{1}{\tau} \Delta f(y, v, \tau) = \lambda f^w(x, v), \end{aligned}$$

which shows that $f^w(x, \cdot)$ is positively homogeneous. We prove now that $f^w(x, \cdot)$ is subadditive. Let $v_1, v_2 \in X$. Fix $k > 0$ and $t > 0$. Then

$$\frac{1}{t} \Delta f(y, v_1 + v_2, t) = \frac{1}{t} \Delta f(y, v_1, t) + \frac{1}{t} \Delta f(y + tv_1, v_2, t).$$

Put $z = y + tv_1$, $\kappa = k + \|v_1\|$. Now $k \rightarrow \infty$ is equivalent to $\kappa \rightarrow \infty$, and if $y \in x + ktB$, then

$$z = y + tv_1 \in x + (k + \|v_1\|)tB = x + \kappa tB.$$

Therefore

$$\begin{aligned} f^w(x, v_1 + v_2) &= \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, v_1 + v_2, t) \\ &\leq \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, v_1, t) + \lim_{\kappa \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{z \in x + \kappa tB} \frac{1}{t} \Delta f(z, v_2, t) \\ &= f^w(x, v_1) + f^w(x, v_2). \end{aligned}$$

b) Let f be Lipschitz of range K near x and $v_1, v_2 \in X$. Then for sufficiently small $t > 0$ we have

$$\begin{aligned} \frac{1}{t} \Delta f(y, v_1, t) &= \frac{1}{t} \Delta f(y, v_2, t) + \frac{1}{t} (f(y + tv_1) - f(y + tv_2)) \\ &\leq \frac{1}{t} \Delta f(y, v_2, t) + K \|v_1 - v_2\|, \end{aligned}$$

whence we obtain easily

$$f^w(x, v_1) \leq f^w(x, v_2) + K\|v_1 - v_2\|.$$

Interchanging v_1 and v_2 we get

$$\|f^w(x, v_1) - f^w(x, v_2)\| \leq K\|v_1 - v_2\|.$$

c) In fact formula (2) takes \limsup in $t \rightarrow 0^+$, $y \rightarrow x$ belonging to a narrower set in comparison to formula (1), whence (3) follows immediately.

d) We have

$$\frac{1}{t} \Delta f(y, -v, t) = -\frac{1}{t} \Delta f(y - tv, v, t) = \frac{1}{t} \Delta(-f)(z, -v, t),$$

where $z = y - tv$. Put $\kappa_1 = k - \|v\|$, $\kappa_2 = k + \|v\|$. We have then

$$\begin{aligned} (-f)^w(x, v) &= \lim_{\kappa_1 \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{z \in x + \kappa_1 t B} \frac{1}{t} \Delta(-f)(z, v, t) \\ &\leq \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, -v, t) = f^w(x, -v) \\ &\leq \lim_{\kappa_2 \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{z \in x + \kappa_2 t B} \frac{1}{t} \Delta(-f)(z, v, t) = (-f)^w(x, v). \end{aligned}$$

This chain of inequalities proves d). □

Further we use the notations:

$$L^* f^w(x) = \sup \{f^w(x, v) \mid v \in B\}, \quad L_* f^w(x) = \inf \{f^w(x, v) \mid v \in B\}.$$

Proposition 2. a) It holds

$$-L^* f^w(x) \leq L_* f^w(x) \leq L^* f^w(x).$$

b) The function $f^w(x, \cdot)$ is Lipschitz if and only if $L^* f^w(x) < \infty$, in which case $L^* f^w(x)$ is the minimal Lipschitz constant for $f^w(x, \cdot)$.

Proof. a) We have

$$\begin{aligned} -L^* f^w(x) &= -\sup \{f^w(x, -v) \mid v \in B\} = -\sup \{(-f)^w(x, v) \mid v \in B\} \\ &= -\sup \left\{ \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta(-f)(y, v, t) \mid v \in B \right\} \\ &= \inf \left\{ \lim_{k \rightarrow \infty} \liminf_{t \downarrow 0} \inf_{y \in x + ktB} \frac{1}{t} \Delta f(y, v, t) \mid v \in B \right\} \\ &\leq \inf \left\{ \lim_{k \rightarrow \infty} \limsup_{t \downarrow 0} \sup_{y \in x + ktB} \frac{1}{t} \Delta f(y, v, t) \mid v \in B \right\} \\ &= \inf \{f^w(x, v) \mid v \in B\} = L_* f^w(x) \leq \sup \{f^w(x, v) \mid v \in B\} = L^* f^w(x). \end{aligned}$$

b) Obviously, if $L^* f^w(x) = \infty$, then the function $f^w(x, \cdot)$ is not Lipschitz. Let $L^* f^w(x) < \infty$. It is clear that $f^w(x, \cdot)$ cannot be Lipschitz of range less than $L^* f^w(x)$. For arbitrary $v, v_0 \in X$ we have

$$f^w(x, v) = f^w(x, (v - v_0) + v_0) \leq f^w(x, v - v_0) + f^w(x, v_0) \leq L^* f^w(x) \|v - v_0\| + f^w(x, v_0).$$

Interchanging v and v_0 we get

$$\|f^w(x, v) - f^w(x, v_0)\| \leq L^* f^w(x) \|v - v_0\|.$$

□

Corollary 1. *If $f : X \rightarrow \overline{\mathbb{R}}$ is Lipschitz with constant K near x , then $L^* f^w(x) \leq K$.*

Proof. According to Proposition 1 b) the function $f^w(x, \cdot)$ is Lipschitz of range K and therefore $L^* f^w(x) \leq K$. □

The following example shows that the property b) from Proposition 2 cannot be reverted.

Example 2. *The function $f : \mathbb{R} \rightarrow \mathbb{R}$,*

$$f(x) = \begin{cases} x^{3/2} \sin(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is not Lipschitz near $x = 0$, but it possesses a Lipschitz weakened derivative $f^w(0, v) = 0$.

To show that the function f in this example is not Lipschitz near $x = 0$ put

$$x_n = \frac{1}{(2n - 3/2)\pi}, \quad y_n = \frac{1}{2n\pi}.$$

Then $x_n \rightarrow 0, y_n \rightarrow 0$ and

$$\frac{f(x_n) - f(y_n)}{x_n - y_n} = \frac{4n}{3\pi^{1/2}(2n - 3/2)^{1/2}} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

We conclude this section with an example showing that the finiteness of $L_* f^w(x)$ does not imply the finiteness of $L^* f^w(x)$ and consequently the Lipschitz property of $f^w(x, \cdot)$.

Example 3. *The function $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = |x|^{1/2}$, satisfies*

$$f^w(0, v) = \begin{cases} +\infty, & v \neq 0, \\ 0, & v = 0. \end{cases}$$

Therefore $L_ f^w(0) = 0$ and $L^* f^w(0) = 0$.*

3 The weakened subdifferential

Let $f : X \rightarrow \overline{\mathbb{R}}$, $x \in \text{int dom } f$. We introduce the weakened subdifferential $\partial^w f(x)$ of f at x as

$$\partial^w f(x) = \{\zeta \in X^* \mid \langle \zeta, v \rangle \leq f^w(x, v) \text{ for all } v \in X\}.$$

We denote by $\|\cdot\|_*$ the norm in X^* , that is $\|\zeta\|_* = \sup \{\langle \zeta, v \rangle \mid v \in X, \|v\| \leq 1\}$, and by B_* the open unit ball in X^* . The weakened gradient obeys the following properties, which we give without proof.

Proposition 3. a) $\partial^w f(x)$ is convex and weak*-closed subset of X^* .

b) $\partial^w f(x)$ is nonempty if and only if

$$L_* f^w(x) > -\infty. \quad (4)$$

If this condition is satisfied, then there exists at least one element $\zeta \in \partial^w f(x)$, such that $\|\zeta\|_* \leq -L_* f^w(x)$. We have in this case

$$\overline{f^w}(x, v) = \sup \{\langle \zeta, v \rangle \mid \zeta \in \partial^w f(x)\}, \quad (5)$$

where $\overline{f^w}(x, \cdot)$ is the closed hull of the convex function $f^w(x, \cdot)$.

c) If $L^* f^w(x) < \infty$, then $\partial^w f(x)$ is nonempty weak*-compact convex subset of X^* and

$$f^w(x, v) = \sup \{\langle \zeta, v \rangle \mid \zeta \in \partial^w f(x)\}. \quad (6)$$

In particular

$$L^* f^w(x) = \sup \{\|\zeta\|_* \mid \zeta \in \partial^w f(x)\}. \quad (7)$$

d) $\partial^w f(x) \subset \partial^\circ f(x)$.

Proof. a) Let $\zeta_1, \zeta_2 \in \partial^w f(x)$ and $0 \leq \lambda \leq 1$. Then the vector $\zeta = (1 - \lambda)\zeta_1 + \lambda\zeta_2$ satisfies for arbitrary $v \in X$ the inequality

$$\langle \zeta, v \rangle = (1 - \lambda)\langle \zeta_1, v \rangle + \lambda\langle \zeta_2, v \rangle \leq (1 - \lambda)f^w(x, v) + \lambda f^w(x, v) = f^w(x, v).$$

Hence $\zeta \in \partial^w f(x)$ which shows that $\partial^w f(x)$ is convex.

If $\zeta_n \in \partial^w f(x)$ converges weakly* to $\zeta \in X^*$ and $v \in X$, then

$$\langle \zeta, v \rangle = \lim_{n \rightarrow \infty} \langle \zeta_n, v \rangle \leq f^w(x, v).$$

Therefore $\zeta \in \partial^w f(x)$ and $\partial^w f(x)$ is weakly* closed.

b) Let condition (4) be not satisfied. Suppose that $\zeta \in \partial^w f(x)$ and let $v_0 \in B$ be such that $f^w(x, v_0) < -\|\zeta\|_*$. Then the following chain of inequalities must be find

$$f^w(x, v_0) < -\|\zeta\|_* \leq -\|\zeta\| \|v_0\| \leq -|\langle \zeta, v_0 \rangle| \leq \langle \zeta, v_0 \rangle \leq f^w(x, v_0),$$

a contradiction. Therefore $\partial^w f(x) = \emptyset$ in the case $L_* f^w(x) = -\infty$.

Suppose that $L = L_* f^w(x) > -\infty$. Consider the sets $A_1 = \{(v, r) \in X \times \mathbb{R} \mid r < L\|v\|\}$ and $A_2 = \{(v, r) \in X \times \mathbb{R} \mid r \geq f^w(x, v)\}$. The sets A_1 and A_2 are convex cones in $X \times \mathbb{R}$. To check this property remind that $L \leq f^w(x, 0) = 0$, hence both $-L\|v\|$ and $f^w(x, v)$ are sublinear in v .

The sets A_1 and A_2 do not intersect. Indeed, for each $(v, r) \in A_2, v \neq 0$, it holds

$$r \geq f^w(x, v) = \|v\| f^w(x, v/\|v\|) \geq \|v\| L,$$

and hence $(v, r) \notin A_1$.

The set A_1 is obviously open in $X \times \mathbb{R}$.

According to the Separation Theorem (see e. g. [6]) there exists a hyperplane

$$H : \langle -\zeta, v \rangle + \alpha r = \beta, \quad \zeta \in X^*, \quad \alpha, \beta \in \mathbb{R},$$

separating A_1 and A_2 .

The set A_1 does not admit a vertical separating hyperplane, therefore $\alpha \neq 0$. Without restriction assume that $\alpha = 1$. Since $(0, r)$ belongs to A_1 for $r < 0$ and to A_2 for $r > 0$, we see that $\beta = 0$. Therefore $H : r = \langle \zeta, v \rangle$. We obtain from here that

$$r \leq \langle \zeta, v \rangle \text{ for } (v, r) \in \overline{A_1}, \quad r \geq \langle \zeta, v \rangle \text{ for } (v, r) \in A_2.$$

Since for all $v \in X$ it holds $(v, f^w(x, v)) \in A_2$, we get $f^w(x, v) \geq \langle \zeta, v \rangle, v \in X$, which shows that $\zeta \in \partial^w f(x)$.

On the other hand $(v, L\|v\|) \in \overline{A_1}$, whence $L\|v\| \leq \langle \zeta, v \rangle$ and $\langle \zeta, -v \rangle \leq -L\|v\|$ for $v \in X$. This inequality shows that $\|\zeta\| \leq -L$.

The function $f^w(x, \cdot)$ is convex. It is simple to check using the sublinearity of $f^w(x, \cdot)$ that for its conjugate

$$(f^w)^*(x, \zeta) = \sup_{v \in X} \{\langle \zeta, v \rangle - f^w(x, v)\}$$

it holds

$$(f^w)^*(x, \zeta) = \delta(\zeta \mid \partial^w f(x)),$$

where $\delta(\cdot \mid \partial^w f(x))$ denotes the indicator of $\partial^w f(x)$. Therefore

$$\delta^*(v \mid \partial^w f(x)) = \sup \{\langle \zeta, v \rangle \mid \zeta \in \partial^w f(x)\} = (f^w)^{**}(x, v) = (\bar{f})^w(x, v).$$

c) In Proposition 2 we showed that $f^w(x, \cdot)$ is Lipschitz of range $L^* f^w(x)$. In particular $f^w(x, \cdot)$ is continuous and therefore $(\bar{f})^w(x, v) = f^w(x, v)$, whence formula (6) follows from (5). Indeed, we have

$$\begin{aligned} L^* f^*(x) &= \sup_{v \in B} f^w(x, v) = \sup_{v \in B} \sup_{\zeta \in \partial^w f(x)} f^w(x, v) \\ &= \sup_{\zeta \in \partial^w f(x)} \sup_{v \in B} f^w(x, v) = \sup \{\|\zeta\|_* \mid \zeta \in \partial^w f(x)\}, \end{aligned}$$

which proves (6).

The weak* compactness of $\partial^w f(x)$ follows from the Alaoglu Theorem (see Rudin [5]).

d) Let $\zeta \in \partial^w f(x)$. If $v \in X$, then

$$\langle \zeta, v \rangle \leq f^w(x, v) \leq f^\circ(x, v).$$

Therefore $\zeta \in \partial^\circ f(x)$. □

Introducing

$$L^* f^\circ(x) = \sup \{f^\circ(x, v) \mid v \in B\}, \quad L_* f^\circ(x) = \inf \{f^\circ(x, v) \mid v \in B\},$$

we can reformulate Proposition 3 for the Clarke gradient of not necessarily Lipschitz near x functions. In particular $\partial^\circ f(x)$ is empty if and only if $L_* f^\circ(x) = -\infty$. The inequality $f^w(x, v) \leq f^\circ(x, v)$ implies

$$L_* f^w(x) \leq L_* f^\circ(x) \quad \text{and} \quad L^* f^w(x) \leq L^* f^\circ(x).$$

In particular $L_* f^\circ(x) = -\infty$ implies $L_* f^w(x) = -\infty$.

In connection with assertion b) of the above proposition, the next example shows that the weakened subdifferential can be empty.

Example 4. The function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^{1/3}$, has at $x = 0$ weakened derivatives

$$f^w(0, v) = \begin{cases} +\infty, & v > 0, \\ 0, & v = 0, \\ -\infty, & v < 0, \end{cases}$$

whence $L_* f^w(x) = -\infty$ and $\partial^w f(x) = \emptyset$.

4 Differentiability

In this section we investigate the case, when $\partial^w f(x)$ consists of a single point. To characterize this situation we introduce the following definition. We say that the function $f : X \rightarrow \overline{\mathbb{R}}$ has a weakened differential $D_w f(x) \in X^*$ at the point $x \in \text{int dom } f$, if $\partial^w f(x)$ is the singleton $\{D_w f(x)\}$.

Here is our main result.

Theorem 2. Let f be weakened differentiable at x . Then

a) $f^w(x, v) = \langle D_w f(x), v \rangle$ for each $v \in X$. In particular $f^w(x, \cdot)$ is Lipschitz with constant $\|D_w f(x)\|$.

b) The Gâteaux derivative $D_G f(x)$ exists and $D_w f(x) = D_G f(x)$.

Proof. a) Obviously

$$f^w(x, v) = \max \{ \langle \zeta, v \rangle \mid \zeta \in \partial^w f(x) \} = \langle D_w f(x), v \rangle.$$

b) Recall that f is said to possess a Gâteaux differential $D_G f(x) \in X^*$, if for all $v \in X$ it holds $\langle D_G f(x), v \rangle = \lim_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t)$. The Gâteaux differentiability of f at x and the equality $D_G f(x) = D_w f(x)$ follows from the equality $\langle D_w f(x), v \rangle = \lim_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t)$, which in turn follows by the following chain.

$$\begin{aligned} \langle D_w f(x), v \rangle &= -\langle D_w f(x), -v \rangle = -f^w(x, -v) = -(-f)^w(x, v) \\ &\leq -\limsup_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t) = \liminf_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t) \\ &\leq \limsup_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t) = \langle D_w f(x), v \rangle. \end{aligned}$$

This chain of inequalities shows that

$$\langle D_w f(x), v \rangle = \lim_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t), \quad (8)$$

which gives that f is Gâteaux differentiable and $D_w f(x) = D_G f(x)$. \square

Remark 1. *The following assertion is true: For each real $k > 0$ it holds*

$$\langle D_w f(x), v \rangle = \lim_{t \downarrow 0, y \in x + ktB} \frac{1}{t} \Delta f(y, v, t).$$

The proof can be obtained in the same way as the proof of Theorem 2 b.

Recall that $f : X \rightarrow \overline{\mathbb{R}}$ is said to possess an Hadamard differential $D_H f(x) \in X^*$ at $x \in \text{int dom } f$, if for all $v \in X$ it holds $\langle D_H f(x), v \rangle = \lim_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t)$ uniformly in $v \in V$, where V is an arbitrary compact set in X . Obviously, when X is finite-dimensional, then f is Hadamard differentiable at x if and only if f is Fréchet differentiable at x .

The following theorem strengthens Theorem 2 for Lipschitz functions and is proved by similar estimations.

Theorem 3. *If $f : X \rightarrow \overline{\mathbb{R}}$ is weakened differentiable and Lipschitz near x , then it is also Hadamard differentiable and $D_w f(x) = D_H f(x)$. If in addition X is finite dimensional, then there exists the Fréchet differential $D_F f(x)$ and $D_w f(x) = D_F f(x)$.*

Proof. Let $V \subset X$ be arbitrary compact set. We must show that (8) holds uniformly in $v \in X$. Put for brevity $\zeta = D_w f(x)$ and let $f(x)$ be Lipschitz near x of range K . Fix $v_0 \in X$ and let v be another point of X . Then

$$\begin{aligned} \left| \frac{1}{t} \Delta f(x, v, t) - \langle \zeta, v \rangle \right| &= \left| \frac{1}{t} \Delta f(x, v_0, t) - \langle \zeta, v_0 \rangle \right. \\ &\quad \left. + \frac{1}{t} (f(x + tv) - f(x + tv_0)) - \langle \zeta, v - v_0 \rangle \right| \\ &\leq \left| \frac{1}{t} \Delta f(x, v_0, t) - \langle \zeta, v_0 \rangle \right| + K \|v - v_0\| + \|\zeta\|_* \|v - v_0\|. \end{aligned}$$

Let $\varepsilon > 0$. Then there exists $\delta = \delta(\varepsilon, v) > 0$, such that

$$\left| \frac{1}{t} \Delta f(x, v_0, t) - \langle \zeta, v_0 \rangle \right| < \frac{1}{2} \varepsilon \quad \text{for } 0 < t < \delta.$$

Let $v \in U(v_0) := v_0 + \frac{\varepsilon}{2(K+\|\zeta\|_*)} B$. Then we obtain

$$\left| \frac{1}{t} \Delta f(x, v, t) - \langle \zeta, v \rangle \right| < \varepsilon. \quad (9)$$

Let $V \subset U(v_1) \cup \dots \cup U(v_n)$ and $\delta = \min\{\delta(\varepsilon/2, v_i) \mid i = 1, \dots, n\}$. Inequality (9) is satisfied for $0 < t < \delta$ and arbitrary $v \in V$. Therefore $\zeta = D_H f(x)$.

Since for finite-dimensional spaces the concept of Fréchet and Hadamard differentials coincide, we have in this case $\zeta = D_F f(x)$. \square

We put an open question: *Is an Hadamard differentiable and Lipschitz near x function f also weakened differentiable?*

The following example shows, that the final part of Theorem 3 is not true for infinite dimensional Banach spaces.

Example 5. Let $X = \ell^1$, i.e. the elements of X are sequences of reals $x = (x_1, x_2, \dots)$, for which $\|x\| := \sum_{i=1}^{\infty} |x_i| < \infty$. Put $\varphi_i : \mathbb{R} \rightarrow \mathbb{R}$, $\varphi_i(t) = |t|^{1+1/i}$, $i = 1, 2, \dots$. Define the function $f : X \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} \sum_{i=1}^{\infty} \frac{1}{1+i} \varphi_i(x_i), & \|x\| \leq 1, \\ +\infty, & \|x\| \geq 1. \end{cases}$$

Then f is Lipschitz with constant 1 near $x = 0$. The weakened subdifferential $\partial^w f(x)$ is the singleton $\{0\}$, whence $D_w f(0) = D_H f(0) = 0$. The Fréchet derivative $D_F f(0)$ however does not exist.

The following reasonings explain the example.

Obviously for $x \in B$ we have $0 \leq \varphi_i(x_i) \leq |x_i|$ and $f(x)$ is the sum of a series with non negative members majorized by $\sum_{i=1}^{\infty} |x_i| = \|x\| < \infty$. Hence f is finite on B and $0 \in \text{int dom } f$.

Let $x, y \in B$. Using the Mean Value Theorem we see that

$$f(x) - f(y) = \sum_{i=1}^{\infty} \frac{i}{1+i} (\varphi_i(x_i) - \varphi_i(y_i)) \leq \sum_{i=1}^{\infty} \xi_i |x_i - y_i|,$$

where

$$\xi_i = \text{sign}(|y_i| + \theta_i(|x_i| - |y_i|)) \left(|y_i| + \theta_i(|x_i| - |y_i|) \right), \quad 0 < \theta_i < 1.$$

It is easy to get now $|\xi_i| \leq 1$ and consequently the Lipschitz condition of range 1

$$|f(x) - f(y)| \leq \sum_{i=1}^{\infty} |x_i - y_i| = \|x - y\|, \quad x, y \in B.$$

Now we prove that $f^w(0, v) = 0$. Let $y = t\ell$, $\ell \in X$, $\|\ell\| \leq k$ where $k > 1$. Then again using the Mean Value Theorem we obtain as above

$$\left| \frac{1}{t} \Delta f(y, v, t) \right| \leq (k + \|v\|) \sum_{i=1}^{\infty} |v_i| |t|^{1/i}.$$

Let $\varepsilon > 0$. Choose the integer m so that

$$\sum_{i=m+1}^{\infty} |v_i| < \frac{\varepsilon}{2(k + \|v\|)}.$$

Then for $y \in ktB$, $0 < t < 1$, we have

$$\begin{aligned} \left| \frac{1}{t} \Delta f(y, v, t) \right| &\leq (k + \|v\|) \sum_{i=1}^m |v_i| |t|^{1/i} + \frac{1}{2} \varepsilon \\ &\leq (k + \|v\|) \|v\| |t|^{1/m} + \frac{1}{2} \varepsilon < \varepsilon \end{aligned}$$

for $|t|^{1/m} < (\varepsilon / (2(k + \|v\|) \|v\|))$. Hence $f^w(0, v) = 0$ and $\partial^w f(x) = \{0\}$.

The Fréchet derivative $D_F f(0)$, if exists, should be zero, since the Gâteaux derivative is zero. Then

$$|f(x)| = o(\|x\|) \quad \text{as } x \rightarrow 0.$$

However this relation is not true. Indeed, if $e^i = (0, \dots, 1, 0, \dots)$ (the only unit is on i -th place), then

$$\frac{f(te^i)}{|t|} = \frac{i}{1+i} |t|^{1/i}$$

does not tend uniformly in i to zero as $t \rightarrow 0^+$, a contradiction.

Thus for a locally Lipschitz weakened differentiable function the Fréchet derivative does not necessary exist. According to Theorem 3 this is not the case when X is finite dimensional. The next theorem shows that the Fréchet derivative for a weakened differentiable function on a finite dimensional space exists even without the the hypothesis that f is Lipschitz near x .

Theorem 4. *Let X be finite-dimensional. If $f : X \rightarrow \overline{\mathbb{R}}$ is weakened differentiable, then it is also Fréchet differential and $D_F f(x) = D_w f(x)$.*

Proof. Suppose that $\dim X = n$ and let $\{e^1, \dots, e^n\}$ be a basis of unit vectors. Let

$$\Gamma = \{\gamma = (\gamma_1, \dots, \gamma_n) \mid \gamma_i = \pm 1, i = 1, \dots, n\}.$$

For each $\gamma \in \Gamma$ we put $E_\gamma \text{co} \{\gamma_1 e^1, \dots, \gamma_n e^n\}$ and let $d_\gamma = \text{dist}(0, E_\gamma)$. Obviously $d_\gamma > 0$. This is true, since E_γ is compact and $0 \notin E_\gamma$ (otherwise the vectors $e^1, \dots, \gamma_n e^n$ would be linearly dependent).

Put $d = \min\{d_\gamma \mid \gamma \in \Gamma\}$. Let $v \in B$. Then for some $\gamma \in \Gamma$ we have

$$v = \sum_{i=1}^n v_i e^i = \sum_{i=1}^n |v_i| \gamma_i e^i = \sum_{i=1}^n |v_i| \gamma_i e^i = \sum_{i=1}^n |v_i| \gamma_i e^i, \quad e \in E_\gamma.$$

Therefore

$$1 \geq \|v\| = \sum_{j=1}^n |v_j| \|e\| \geq d \sum_{j=1}^n |v_j| \geq d \|v\|.$$

Thus for each $v \in B$, there exists $\gamma \in \Gamma$, such that

$$v = \sum_{i=1}^n \alpha_i \gamma_i e^i, \quad 0 \leq \alpha_i \leq 1/d. \quad (10)$$

Suppose that f is weakened differentiable at x and $\zeta = D_w f(x)$. Take $\varepsilon > 0$ and choose $\delta > 0$ such that

$$\left| \frac{1}{s} \Delta f(y, p, s) - \langle \zeta, p \rangle \right| < \frac{\varepsilon}{n} d \quad (11)$$

for each $p = \pm e^i$, $i = 1, 2, \dots, n$, $0 < s \leq \delta$, $y \in x + (n-1)sB$. Choose $v \in B$, and let v be represented as in (10). Put $\alpha = \alpha_1 + \dots + \alpha_n$. Let $\bar{\alpha}_1 \leq \dots \leq \bar{\alpha}_n$ be a permutation of $(\alpha_1, \dots, \alpha_n)$. Then

$$\left| \frac{1}{t} \Delta f(x, v, t) - \langle \zeta, v \rangle \right| \leq \sum_{j=1}^n \left| \frac{1}{t} \Delta f(y_{j-1}, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j) \right|, \quad (12)$$

where

$$y_0 = x, \quad y_j = x + t \sum_{i=1}^j \bar{\alpha}_i \bar{\gamma}_i \bar{e}^i, \quad j = 1, \dots, n, .$$

(here $\bar{\gamma}_j$ and \bar{e}^j denote the corresponding to $\bar{\alpha}_j$ permutation of γ_j and e^j). We have for $0 < t < \delta d$

$$\left| \frac{1}{t} \Delta f(y_{j-1}, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j, t) - \langle \zeta, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j \rangle \right| < \frac{\varepsilon}{n}. \quad (13)$$

If $\bar{\alpha}_j = 0$ this inequality is obvious. If $\bar{\alpha}_j > 0$ it follows from (11) with $s = \bar{\alpha}_j t$ and $y = y_j$. Indeed, we have

$$0 < s = \bar{\alpha}_j t \leq \frac{\delta}{d} d = \delta, \\ y_{j-1} = x + t \sum_{i=1}^{j-1} \bar{\alpha}_i \bar{\gamma}_i \bar{e}^i \in x + \bar{\alpha}_j t (n-1)B = x + (n-1)sB$$

(here we use that $\bar{\alpha}_i$ is a permutation of α_i) and

$$\begin{aligned} & \left| \frac{1}{t} \Delta f(y_{j-1}, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j, t) - \langle \zeta, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j \rangle \right| \\ &= \bar{\alpha}_j \left| \frac{1}{\bar{\alpha}_j t} \Delta f(y_{j-1}, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j, t) - \langle \zeta, \bar{\alpha}_j \bar{\gamma}_j \bar{e}^j \rangle \right| \leq \bar{\alpha}_j \frac{\varepsilon}{n} d \leq \frac{\varepsilon}{n}. \end{aligned}$$

Inequality (13) applied to (12) gives

$$\left\| \frac{1}{t} \Delta f(x, v, t) - \langle \zeta, v \rangle \right\| \leq n \frac{\varepsilon}{n} = \varepsilon$$

for all $v \in B$ and $0 < t < \delta d$, which shows that f is Fréchet differentiable and $D_F f(x) = D_G f(x)$. \square

In general the Fréchet differentiability is stronger than the weakened differentiability.

Theorem 5. *If f is Fréchet differentiable at x , then it is also weakened differentiable at x and $D_w f(x) = D_F f(x)$.*

Proof. Fix $v \in X$ and $k > \|v\|$. Put $y = x + ty_1$ with $\|y_1\| < k$, and $\zeta = D_F f(x)$. Then

$$\begin{aligned} & \left| \frac{1}{t} \Delta(y, v, t) - \langle \zeta, v \rangle \right| \\ & \leq \left| \frac{1}{t} \Delta(y, v + y_1, t) - \langle \zeta, v + y_1 \rangle \right| + \left| \frac{1}{t} \Delta(x, y_1, t) - \langle \zeta, y_1 \rangle \right|. \end{aligned} \quad (14)$$

Fix $\varepsilon > 0$. The Fréchet differentiability of f gives that there exists $\delta > 0$, such that for each $z \in (\|v\| + k)B$ and $0 < t < \delta$ we have

$$\left| \frac{1}{t} \Delta(x, z, t) - \langle \zeta, z \rangle \right| < \frac{1}{2} \varepsilon.$$

Since both y_1 and $v + y_1$ are in $(\|v\| + k)B$, we get from (14).

$$\left| \frac{1}{t} \Delta(y, v, t) - \langle \zeta, v \rangle \right| \leq \varepsilon.$$

This observation shows that $f^w(x, v) = \langle \zeta, v \rangle$ and consequently $D_F f(x) = D_w f(x)$. \square

Combining Theorems 4 and 5 we get the following result, formulated in a similar manner to Theorem 1.

Theorem 6. *Let X be finite-dimensional and let $f : X \rightarrow \overline{\mathbb{R}}$, $x \in \text{int dom } f$ and $\zeta \in X^*$. Then the following assertions are equivalent:*

- a) f is Fréchet differentiable at x and $D_F f(x) = \zeta$,
- b) f is Lipschitz near x and $\partial^w f(x)$ is the singleton $\{\zeta\}$.

5 The regularity condition

In the next section we introduce some basic calculus rules for the weakened gradient. In some of these rules the following regularity condition is important. We say that f is weakened regular (or briefly w -regular) at x if the next three requirements are satisfied:

- i) $L^* f^w(x) < \infty$,
- ii) For all $v \in X$ there exists the usual one-sided directional derivative

$$f'(x, v) = \lim_{t \downarrow 0} \frac{1}{t} \Delta f(x, v, t),$$

- iii) For all $v \in X$ it holds $f'(x, v) = f^w(x, v)$.

Replacing in iii) $f^w(x, v)$ by $f^\circ(x, v)$ we obtain the notion of Clarke regularity. The inequalities $f'(x, v) \leq f^w(x, v) \leq f^\circ(x, v)$ show that the weakened regularity is weaker than the Clarke regularity. Here there are some conditions that guarantee weakened regularity.

Proposition 4. a) If f is weakened differentiable (in particular Fréchet differentiable), then it is w -regular at x .

b) If f admits a Gâteaux differential $D_G f(x)$ and if it is w -regular at x , then it is weakened differentiable at x and $D_G f(x) = D_w f(x)$.

c) A finite linear combination with nonnegative scalars of functions w -regular at x is w -regular at x .

Proof. a) Let $\zeta = D_w f(x)$. Then $f^w(x, v) = \langle \zeta, v \rangle$. The definition of the weakened derivative gives that for each $\varepsilon > 0$ there exists $\delta = \delta(x, v, \varepsilon) > 0$, such that the inequality

$$\left| \frac{1}{t} \Delta f(x, v, t) - \langle \zeta, v \rangle \right| < \varepsilon$$

holds for $0 < t < \delta$. Therefore the directional derivative $f'(x, v)$ exists and $f'(x, v) = \langle \zeta, v \rangle = f^w(x, v)$.

b) Put $\zeta = D_G f(x)$. Then $\langle \zeta, v \rangle = f'(x, v) = f^w(x, v)$, whence we see that f is weakened differentiable at x and $D_w f(x) = \zeta$.

c) The general case follows by obvious induction from the case of two functions. Since

$$(s f)'(x, v) = s f'(x, v), \quad (s f)^w(x, v) = s f^w(x, v),$$

for $s \geq 0$, it suffices to prove that

$$(f^1 + f^2)' = (f^1 + f^2)^w. \quad (15)$$

The existence of $(f^1 + f^2)'$ is evident and

$$(f^1 + f^2)' = (f^1)' + (f^2)' = (f^1)^w + (f^2)^w \geq (f^1 + f^2)^w.$$

The last inequality is clear from the definition of the weakened directional derivative. Since the opposite inequality is always true, we obtain (15). \square

6 Calculus rules

In this section we give formulas facilitating the calculation of $\partial^w f(x)$, when f is build up from simple functionals through linear combination, by maximization, composition and so on. In all these rules the condition

$$L^* f^w(x) < \infty \quad (16)$$

occurs to play an important role. As we saw in Proposition 2 this is satisfied if and only if $f^w(x, v)$ is Lipschitz in v , which in turn is weakening of the condition f to be Lipschitz near x . For brevity we omit the proofs of the calculus rules.

Proposition 5 (Scalar Multipliers). *If the function f satisfies (16), then for any real s it holds*

$$\partial^w (s f)(x) = s \partial^w f(x). \quad (17)$$

Proof. If $s \geq 0$, then (17) follows from

$$(sf)^w(x, v) = sf^w(x, v).$$

It suffices now to prove (17) for $s = -1$. In this case $\zeta \in \partial^w(-f)(x)$ is equivalent to

$$\langle \zeta, v \rangle = \langle -\zeta, -v \rangle \leq (-f)^w(x, -v), \quad v \in X,$$

and consequently to $\zeta \in -\partial^w f(x)$. □

Proposition 6 (Local Extrema). *If the function f satisfies (16) and attains a local extremum at x , then*

$$0 \in \partial^w f(x). \quad (18)$$

Proof. Since $\partial^w(-f)(x) = -\partial^w f(x)$ it suffices to consider the case of a local minimum. Then

$$0 \leq f'(x, v) \leq f^w(x, v), \quad v \in X,$$

whence (18) follows immediately. □

Proposition 7 (Finite Sums). *If $L_* f_i^w(x) < \infty$, $i = 1, 2, \dots, n$, then*

$$\partial^w \left(\sum_{i=1}^n s_i f_i(x) \right) \subset \sum_{i=1}^n s_i \partial^w f_i(x). \quad (19)$$

This inclusion turns into an equality in each of the cases:

- i) All but at most one of the functions f_i are weakened differentiable (in particular Fréchet differentiable).*
- ii) All the functions f_i are w -regular and all the scalars are non-negative (or more general, all the functions $\text{sign}_{s_i} f_i$ are w -regular).*

Proof. In view of Proposition 5 and obvious induction argument it suffices to consider the case $n = 2$, $s_1 = s_2 = 1$. We must prove that

$$(f_1 + f_2)^w \leq (f_1)^w + (f_2)^w,$$

which is clear from the definition of the weakened directional derivative.

i) Adding all the weakened differentiable functions together we reduce the proof to the case of two functions $f_1 + f_2$ with f_1 being weakened differentiable and therefore $(f_1)^w(x, v) = \langle D_w f_1(x), v \rangle$. The equality in (19) follows by the equality

$$(f_1 + f_2)^w(x, v) = \langle D_w f_1(x), v \rangle + (f_2^w)(x, v),$$

which easily follows from the definition of the weakened derivative.

ii) Like above we may confine to the sum of two w -regular functions $f_1 + f_2$. The assertion is deduced from the inequalities

$$\begin{aligned} (f_1)^w(x, v) + (f_2)^w(x, v) &= (f_1)'(x, v) + (f_2)'(x, v) \\ &= (f_1 + f_2)'(x, v) \leq (f_1 + f_2)^w(x, v) \leq (f_1)^w(x, v) + (f_2)^w(x, v). \end{aligned}$$

□

We are going to establish a Lebourg type Mean Value Theorem for the weakened subdifferential. Given x and y in X , we denote by $[x, y]$ and (x, y) correspondingly the closed and open line segment with end points x and y . Using the notation $x_t = (1 - t)x + ty$, then

$$[x, y] = \{x_t \mid 0 \leq t \leq 1\}, \quad (x, y) = \{x_t \mid 0 < t < 1\}, \quad .$$

We prove first the following lemma.

Lemma 1. *Let x and y be fixed points in X and let the continuous function f fulfills $L^* f^w(x_t) < \infty$ for each $t \in [0, 1]$. Then the function*

$$g : [0, 1] \rightarrow \mathbb{R}, \quad g(t) = f(x_t)$$

satisfies the inclusion $\partial^w g(t) \subset \langle \partial^w f(x_t), y - x \rangle$.

Proof. In this inclusion the two closed convex sets are in \mathbb{R} and hence they are intervals. Therefore, it suffices to prove that for $v \neq \pm 1$ we have

$$\max \{ \partial^w g(t) v \} \leq \max \{ \langle \partial^w f(x_t), (y - x)v \rangle \} .$$

This follows from

$$\begin{aligned} g^w(t, v) &= \lim_{k \rightarrow \infty} \limsup_{\lambda \downarrow 0} \sup_{s \in t+k\lambda[-1, 1]} \frac{1}{\lambda} \Delta g(s, v, \lambda) \\ &= \lim_{k \rightarrow \infty} \limsup_{\lambda \downarrow 0} \sup_{s \in t+k\lambda[-1, 1]} \frac{1}{\lambda} \Delta f(x_s, v(y - x), \lambda) \\ &= \lim_{k \rightarrow \infty} \limsup_{\lambda \downarrow 0} \sup_{z \in x_t+k\lambda\|y-x\|B} \frac{1}{\lambda} \Delta f(z, v(y - x), \lambda) \\ &= f^w(x_t, v(y - x)) = \max \{ \langle \partial^w f(x_t), (y - x)v \rangle \} . \end{aligned}$$

□

Proposition 8 (Mean Value Theorem). *Let x and y be points in X and let the continuous function f fulfills $L^* f^w(x_t) < +\infty$, where $x_t = (1 - t)x + ty$, for all $t \in [0, 1]$. Then there exists a point u in the segment (x, y) , such that*

$$f(y) - f(x) = \langle \partial^w f(u), y - x \rangle .$$

Proof. Consider the function

$$\theta : [0, 1] \rightarrow \mathbb{R}, \quad \theta(t) = f(x_t) + t(f(x) - f(y)) .$$

This function is continuous on $[0, 1]$ and satisfies $\theta(0) = \theta(1) = f(x)$, so there is a point t in $(0, 1)$ at which θ attains a local minimum or maximum. By Proposition 6 we have $0 \in \partial^w \theta(x)$. Applying Propositions 5 and 6 and Lemma 1 we obtain

$$0 \in \partial^w f(x_t) + (f(x) - f(y)) \subset \langle \partial^w f(u), y - x \rangle + f(x) - f(y) ,$$

where $u = x_t$.

□

Corollary 2. *Let f be continuous on the convex set $V \subset X$ and let $\sup \{L^* f^w(x) \mid x \in V\} = K < \infty$. Then f is Lipschitz with constant K on V .*

Proof. Let $x, y \in V$. Then for some $u \in (x, y)$ we have

$$f(y) - f(x) = \langle \partial^w f(u), y - x \rangle$$

and therefore

$$f(y) - f(x) \leq \sup_{u \in [x, y]} |f^w(u, y - x)| \leq \sup_{u \in [x, y]} L^* f^w(u) \|y - x\| \leq K \|y - x\|.$$

□

The considered calculus rules show much similarities between the weakened and the Clarke generalized gradient both in the formulations and in the proves. We give further two chain rules for the weakened subdifferentials. We skip the proofs, since they are obtained similarly to the analogous chain rules for the Clarke subdifferentials. We use in the formulation the following agreement. The component functions of a function $h : X \rightarrow \mathbb{R}^n$ are denoted by h_1, \dots, h_n . The components of an operator $\zeta \in L(X, \mathbb{R}^n)$ are denoted by ζ_1, \dots, ζ_n . We put

$$\partial^w h(x) = \{\zeta \mid \zeta_i \in (h_i)^w(x), i = 1, \dots, n\}.$$

The space $(\mathbb{R}^n)^*$ is identified with \mathbb{R}^n and an element $\alpha \in \partial^w h(x)$ is an n -dimensional vector $(\alpha_1, \dots, \alpha_n)$.

Proposition 9 (Chain Rule I). *Suppose that the function $h : X \rightarrow \mathbb{R}^n$ is Lipschitz near x and the function $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is Lipschitz near $h(x)$. Then the weakened gradient $\partial^w f(x)$ of the function $f = g \circ h$ is not empty and*

$$\begin{aligned} \partial^w f(x) &\subset \overline{\text{conv}} \{\alpha \circ \zeta \mid \alpha \in \partial^w g(h(x)), \zeta \in \partial^w h(x)\} \\ &= \overline{\text{conv}} \left\{ \sum_{i=1}^n \alpha_i \circ \zeta_i \mid \alpha \in \partial^w g(h(x)), \zeta_i \in \partial^w h_i(x) \right\}, \end{aligned}$$

where $\overline{\text{conv}}$ denotes the weak* closed convex hull and can be replaced by conv if X is finite-dimensional. An equality holds under any one of the following additional hypotheses:

- i) Each h_i is w -regular at x , g is w -regular at $h(x)$ and every element α of $\partial^w g(h(x))$ has nonnegative components.
- ii) The function g is weakened differentiable at $h(x)$ and $n = 1$ (in this case conv is superfluous).
- iii) Each h_i is weakened differentiable at $h(x)$ (and then $\overline{\text{conv}}$ is superfluous).

Proposition 10 (Chain Rule II). *Suppose that the function $h : X \rightarrow \mathbb{R}^n$ is continuous at x and satisfies the condition $\max_{1 \leq i \leq n} L^*(h_i)^w(x) < \infty$, and let the function $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is Lipschitz near $h(x)$. Then the weakened gradient $\partial^w f(x)$ of the composition $f = g \circ h$ is not empty and*

$$\partial^w f(x) \subset \overline{\text{conv}} \{\alpha \circ \zeta \mid \alpha \in \partial^\circ g(h(x)), \zeta \in \partial^w h(x)\},$$

where ∂° stands for the Clarke subdifferential and $\overline{\text{conv}}$ can be replaced by conv if X is finite-dimensional.

The following three propositions are straightforward application of the above chain rules.

Proposition 11 (Pointwise Maxima). *Suppose that the functions f_1, \dots, f_n satisfy the condition $\max_{1 \leq i \leq n} L^*(f_i)^w(x) < \infty$, Define the function $f(x) = \max_{1 \leq i \leq n} f_i(x)$ and let $I(x)$ be the set of indexes for which $f_i(x) = f(x)$. Then $\partial^w f(x)$ is not empty and*

$$\partial^w f(x) \subset \{\partial^w f_i(x) \mid i \in I(x)\}.$$

If the functions f_i , $i \in I(x)$, are w -regular at x , then the inclusion can be replaced by an equality, and the function f is w -regular at x .

Proposition 12 (Products). *Let f_1 and f_2 satisfy*

$$\max(L^* f_1^w(x), L^* f_2^w(x)) < \infty. \quad (20)$$

Then the weakened subdifferential $\partial^w(f_1 f_2)(x)$ is not empty and

$$\partial^w(f_1 f_2)(x) \subset \partial^w f_1(x) f_2(x) + f_1(x) \partial^w f_2(x).$$

If in addition $f_1(x) \geq 0$, $f_2(x) \geq 0$, and if both f_1 and f_2 are w -regular at x , then the inclusion can be replaced by an equality, and $f_1 f_2$ is w -regular at x .

Proposition 13 (Quotients). *Let f_1 and f_2 satisfy (20), and suppose that $f_2(x) \neq 0$. Then the weakened subdifferential $\partial^w(f_1/f_2)(x)$ is not empty and*

$$\partial^w(f_1/f_2)(x) \subset (\partial^w f_1(x) f_2(x) - f_1(x) \partial^w f_2(x)) / f_2^2(x).$$

If in addition $f_1(x) \geq 0$, $f_2(x) > 0$, and if both f_1 and $-f_2$ are w -regular at x , then the inclusion can be replaced by an equality, and f_1/f_2 is w -regular at x .

7 Final Remarks

The initial motivation of this work was applying similar approach to that of Clarke [2] to define a subdifferential of the function $f : X \rightarrow \overline{\mathbb{R}}$, such that the single-valuedness of the subdifferential at a given point x to be equivalent to the Fréchet differentiability of f at x (as it is known the Clarke subdifferential is related to the strict differentiability). We fulfilled this task, defining the notion of the weakened subdifferential, at least in the finite-dimensional case, see Theorem 6. The inclusion $\partial^w f(x) \subset \partial^\circ f(x)$, motivating the name *weakened* for the new subdifferential, shows that the weakened subdifferential $\partial^w f(x)$ can be more sensitive in applications than the Clarke subdifferential $\partial^\circ f(x)$. For instance, consider the optimization problem $g(x) = f(x) + \varepsilon x \rightarrow \text{extr}$, where $f(x)$ is the function from Example 1 and $\varepsilon \neq 0$. Then obviously $x = 0$ is not a solution of this problem. This can be established applying weakened subdifferentials on the base of Proposition 6, since $0 \notin \partial^w g(0) = \{\varepsilon\}$. At the same time, when $-1 \leq \varepsilon \leq 1$, this does not follow by analogous assertions for the Clarke subdifferential, since $0 \in \partial^\circ g(0) = [-1 + \varepsilon, 1 + \varepsilon]$. This observation could serve as an impulse to investigate whether it is possible to strengthen other known results replacing the Clarke subdifferential by the weakened subdifferential. A strong motivation to prefer the weakened subdifferential instead of other existing generalized subdifferentials, e. g. [3], is that the weakened subdifferential, as the previous section demonstrates, preserves the good calculus rules of the Clarke subdifferential.

References

- [1] J. P. Aubin, H. Frankowska: *Set-valued analysis*. Birkhäuser, Boston-Basel-Berlin, 1990.
- [2] F. H. Clarke: *Optimization and nonsmooth analysis*. Classics in Applied Mathematics 5, SIAM, Philadelphia PA, 1990.
- [3] J. P. Penot: *Variations on the theme of nonsmooth analysis: another subdifferential*. In: V. F. Demyanov, D. Pallaschke (eds), *Nondifferentiable optimization: motivations and applications*, Lecture Notes in Econom. and Math. Systems, 255, Springer Verlag, Berlin and New York, 1985.
- [4] B. N. Pshenichnyi, *Convex analysis and extremal problems* (in Russian). Nauka, Moscow, 1980.
- [5] W. Rudin: *Functional analysis*. McGraw-Hill Book Company, New York, 1973.
- [6] H. H. Schaefer: *Topological vector spaces*. The MacMillan Company, New York, 1966.