Dedicated to
Professor Czesław Olech

Control and Cybernetics

vol. **34** (2005) No. 3

Paraconvex analysis

by

Stefan Rolewicz

Institute of Mathematics of the Polish Academy of Sciences
Śniadeckich 8, 00-956 Warszawa 10, P.O.Box 21, Poland
e-mail: rolewicz@impan.gov.pl

Abstract: In the theory of optimization an essential role is played by the differentiability of convex functions. In this paper we shall try to extend the results concerning differentiability to a larger class of functions called strongly $\alpha(\cdot)$-paraconvex.

Let $(X, \|.\|)$ be a real Banach space. Let $f(x)$ be a real valued strongly $\alpha(\cdot)$-paraconvex function defined on an open convex subset $\Omega \subset X$, i.e.

$$f\big(tx+(1-t)y\big) \le tf(x)+(1-t)f(y)+C\min[t,(1-t)]\alpha(\|x-y\|).$$

Then there is a set of the first Baire category $A_F \subset \Omega$ such that the function $f(\cdot)$ is Fréchet differentiable at every point $x_0 \in \Omega \setminus A_F$.

Keywords: $\alpha(\cdot)$-paraconvex functions, Fréchet differentiability.

1. The $\alpha(\cdot)$-paraconvex and strongly $\alpha(\cdot)$-paraconvex functions

Let $(X, \|.\|)$ be a real Banach space. Let $f(x)$ be a real valued convex continuous function defined on an open convex subset $\Omega \subset X$, i.e.

$$f\big(tx+(1-t)y\big) \le tf(x)+(1-t)f(y). \tag{1}$$

We recall that a set $B \subset \Omega$ of second Baire category is called *residual* if its complement $\Omega \setminus B$ is of the first Baire category. Mazur (1933) proved that if X is separable, then there is a residual subset A_G such that on the set A_G the function f is Gateaux differentiable. Asplund (1968) showed that if additionally in the dual space X^* there exists an equivalent locally uniformly rotund norm, then there is a residual subset A_F such that on the set A_F the function f is

Fréchet differentiable. The spaces X such that for the dual space X^* there exists an equivalent locally uniformly rotund norm are now called *Asplund spaces*. It can be shown that each reflexive space and spaces having separable duals are Asplund spaces. Even more, a space X is an Asplund space if and only if each its separable subspace $X_0 \subset X$ has a separable dual (Phelps, 1989).

Let $\alpha(t)$ be a nondecreasing function mapping the interval $[0, +\infty)$ into the interval $[0, +\infty]$ such that

$$\lim_{t \downarrow 0} \frac{\alpha(t)}{t} = 0. \tag{2}$$

Let, as before, $(X, \|.\|)$ be a real Banach space. Let $f(x)$ be a real valued continuous function defined on an open convex subset $\Omega \subset X$. We say that the function $f(\cdot)$ is $\alpha(\cdot)$-*paraconvex* if there is a constant $C > 0$ such that for all $x, y \in \Omega$ and $0 \leq t \leq 1$

$$f\big(tx + (1-t)y\big) \leq tf(x) + (1-t)f(y) + C\alpha(\|x - y\|). \tag{3}$$

For $\alpha(t) = t^2$ this definition was introduced in Rolewicz (1979a) and the t^2-paraconvex functions were called simply paraconvex functions. In Rolewicz (1979b) the notion was extended to the case $\alpha(t) = t^\gamma, 1 \leq \gamma \leq 2$, and the t^γ-paraconvex functions were called γ-paraconvex functions.

We say that the function $f(\cdot)$ is *strongly* $\alpha(\cdot)$-*paraconvex* if there is a constant $C_1 > 0$ such that for all $x, y \in \Omega$ and $0 \leq t \leq 1$

$$f\big(tx + (1-t)y\big) \leq tf(x) + (1-t)f(y) + C_1 \min[t, (1-t)]\alpha(\|x - y\|). \tag{4}$$

Of course every function $f(\cdot)$ strongly $\alpha(\cdot)$-paraconvex is also $\alpha(\cdot)$-paraconvex. The converse is not true and the conditions warranting the fact that each $\alpha(\cdot)$-paraconvex is strongly $\alpha(\cdot)$-paraconvex can be found in Rolewicz (2000). In particular each t^γ-paraconvex function, $1 < \gamma \leq 2$, is strongly t^γ-paraconvex.

If

$$\lim_{t \downarrow 0} \frac{\alpha(t)}{t^2} = 0, \tag{5}$$

then an $\alpha(\cdot)$-paraconvex function is convex. If

$$\limsup_{t \downarrow 0} \frac{\alpha(t)}{t^2} < \infty, \tag{6}$$

then an $\alpha(\cdot)$-paraconvex function is a difference of a convex and a quadratic function (Rolewicz, 1980). If

$$\lim_{t \downarrow 0} \frac{\alpha(t)}{t^2} = +\infty, \tag{7}$$

then the class of $\alpha(\cdot)$-paraconvex functions is larger.

The notion of $\alpha(\cdot)$-paraconvex functions can be treated as a uniformization of the notion of approximate convex functions introduced in the papers of Luc, Ngai and Théra (1999, 2000). We recall that a real-valued function $f(\cdot)$ defined on a convex set $\Omega \subset X$ is called *approximate convex* if for arbitrary $x_0 \in \Omega$ and $\varepsilon > 0$ there is $\delta = \delta(\varepsilon, x_0)$ such that for x, y such that $\|x - x_0\| < \delta$ and $\|y - x_0\| < \delta$ we have

$$f\big(tx + (1-t)y\big) \le tf(x) + (1-t)f(y) + \varepsilon \min[t, (1-t)]\|x - y\|. \tag{8}$$

We say that a real-valued function $f(\cdot)$ defined on a convex set $\Omega \subset X$ is called *uniformly approximate convex* if for arbitrary $\varepsilon > 0$ there is $\delta = \delta(\varepsilon)$ such that for x, y such that $\|x - y\| < \delta$ and (8) holds.

PROPOSITION 1.1 (Rolewicz, 2001b) *Let $(X, \|.\|)$ be a real Banach space. Let $f(x)$ be a real valued continuous function defined on an open convex subset $\Omega \subset X$. Then it is uniformly approximate convex if and only if there is $\alpha(\cdot)$ satisfying (2) such that $f(\cdot)$ is $\alpha(\cdot)$-paraconvex.*

We shall recall now the different notions of directional derivatives.

By the *Dini derivative* of a continuous function $f(x)$ at a point x_0 in a direction h we mean the number

$$d^D f\big|_{x_0}(h) = \liminf_{\substack{t \downarrow 0 \\ u \to h}} \frac{f(x_0 + tu) - f(x_0)}{t}, \tag{9}$$

where the lower limit in formula (9) is taken with respect to any sequence $\{t_n\}$ of positive numbers tending to 0 and to any sequence $\{u_n\}$ tending to h.

By the *lower directional derivative* of a continuous function $f(x)$ at a point x_0 in a direction h we mean the number

$$d^{ld} f\big|_{x_0}(h) = \liminf_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0)}{t}. \tag{10}$$

By the *upper directional derivative* of a continuous function $f(x)$ at a point x_0 in a direction h we mean the number

$$d^{ud} f\big|_{x_0}(h) = \limsup_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0)}{t}. \tag{11}$$

If the lower directional derivative is equal to the upper directional derivative we shall call it simply *directional derivative*.

By the *Clarke derivative* of the function $f(x)$ at a point x_0 in a direction h we mean the number

$$d^{Cl} f\big|_{x_0}(h) = \limsup_{\substack{t \downarrow 0 \\ x \to x_0}} \frac{f(x + th) - f(x)}{t}, \tag{12}$$

where the upper limit in formula (12) is taken with respect to any sequence $\{t_n\}$ of positive numbers tending to 0 and to any sequence $\{x_n\}$ of elements belonging to the domain of the function $f(x)$ (i.e., such that $|f(x_n)| < +\infty$) tending to x_0.

It is easy to see that always

$$d^D f\big|_{x_0}(h) \leq d^{ld} f\big|_{x_0}(h) \leq d^{ud} f\big|_{x_0}(h) \leq d^{Cl} f\big|_{x_0}(h). \tag{13}$$

In the case of strongly $\alpha(\cdot)$-paraconvex functions defined on open convex sets we have equality. The proof is based on the following

Proposition 1.2 (Rolewicz, 2000) *Let Ω be an open convex set in a Banach space X. Let $f(\cdot)$ be a strongly $\alpha(\cdot)$-paraconvex function. Then it is a locally Lipschitz function.*

Basing on Proposition 1.2 we can prove:

Theorem 1.1 (Rolewicz, 2001) *Let Ω be an open convex set in a Banach space X. Let $f(\cdot)$ be a strongly $\alpha(\cdot)$-paraconvex function. Then*

$$d^D f\big|_{x_0}(h) = d^{ld} f\big|_{x_0}(h) = d^{up} f\big|_{x_0}(h) = d^{Cl} f\big|_{x_0}(h). \tag{14}$$

The aim of this paper is to show

Theorem 1.2 (Rolewicz, 2005) *Let Ω be an open convex set of an Asplund space X. Let $f(\cdot)$ be a continuous strongly $\alpha(\cdot)$-paraconvex function. Then there is a residual set $A_F \subset \Omega$ such that the function $f(\cdot)$ is Fréchet differentiable at every point $x_0 \in A_F$.*

At the moment I do not know anything about the possibility of replacing the assumption of strong $\alpha(\cdot)$-paraconvexity by $\alpha(\cdot)$-paraconvexity.

2. Uniform approximative subdifferentiability

For the case of convex functions the proof of theorem similar to Theorem 1.2 consists of two parts

(a) a convex function has a subgradient at each point,

(b) if a function has a subgradient at each point, then there is a residual set $A_F \subset \Omega$ such that the function $f(\cdot)$ is Fréchet differentiable at every point $x_0 \in A_F$.

In the classical situation condition (a) is so trivial, that it is not observed at all. But now we are in a different situation. It is necessary to define "subgradients" and to show a strongly $\alpha(\cdot)$-paraconvex function has a "subgradient" at each point. Moreover the corresponding "subdifferentiability" ought to warrant the step (b).

The natural candidate is so the called approximate subgradient introduced by Ioffe and Mordukhovich (see Ioffe, 1984, 1986, 1989, 1990, 2000; Mordukhovich, 1976, 1980, 1988). Namely a linear functional $x^*(\cdot) \in X^*$ will be called an *approximate subgradient* of the function $f(x)$ at a point x if

$$\liminf_{h\to 0} \frac{\big(f(x+h)-f(x)\big)-x^*(h)}{\|h\|} \geq 0. \tag{15}$$

The set of all approximative subgradients of the function $f(\cdot)$ at a point x shall be called *approximate subdifferential* of the function f at the point x and we shall denote it as in the classical case by $\partial f\big|_x$.

Of course $\partial f\big|_{(\cdot)}$ is a multifunction mapping a domain of $\partial f\big|_{(\cdot)}$ into 2^{X^*}.

If for all $x \in \Omega$ $\partial f\big|_x \neq \emptyset$ we say that the function $f(\cdot)$ is *approximate subdifferentiable*.

Unfortunately, till now we are not able to show that an approximate subdifferentiable function $f(\cdot)$ is Fréchet differentiable on a residual set. However, the uniformization of the notion of approximate subdifferentiability has the requested property.

Observe that (15) holds if and only if there is a non-negative non-decreasing function $\beta_x(\cdot)$ defined on the interval $[0,+\infty)$ and such that $\lim_{u\downarrow 0}\beta_x(u)=0$ and

$$\frac{\big(f(x+h)-f(0)\big)-x^*(h)}{\|h\|} \geq -\beta_x(\|h\|). \tag{16}$$

Indeed, the function

$$\beta_x(s)) = \sup_{\{h:\|h\|\leq s\}} \Big|\frac{\big(f(x+h)-f(x)\big)-x^*(h)}{\|h\|}\Big| \tag{17}$$

has the requested property.

Putting $\alpha_x(u)=u\beta_x(u)$ we can rewrite (16) in the form

$$f(x+h)-f(x) \geq x^*(h)-\alpha_x(\|h\|). \tag{18}$$

Unfortunately $\beta_x(\cdot)$ (and thus $\alpha_x(\cdot)$) can be different in each point and we are not able to use this definition for the problem of differentiation on a residual set. Thus there is an idea of a uniformization of this notion.

Let $\alpha(t)$ be a nondecreasing function mapping the interval $[0,+\infty)$ into the interval $[0,+\infty]$ such that

$$\lim_{t\downarrow 0}\frac{\alpha(t)}{t}=0. \tag{2}$$

Let $f(\cdot)$ be a real-valued function defined on an open set Ω of a Banach space X. Let $x\in X$. A linear functional $x^*\in X^*$ such that

$$f(x+h)-f(x) \geq x^*(h)-\alpha(\|h\|) \tag{19}$$

is called a *uniform approximate subgradient of the function* $f(\cdot)$ *at* x *with the modulus* $\alpha(\cdot)$ (or briefly $\alpha(\cdot)$-*subgradient* of the function $f(\cdot)$ at x). The set of all $\alpha(\cdot)$-subgradients of the function $f(\cdot)$ at x will be called the $\alpha(\cdot)$-*subdifferential of the function* $f(\cdot)$ *at* x and it will be denoted by $\partial_\alpha f\big|_x$.

We say that a function $f(x)$ is $\alpha(\cdot)$-*subdifferentiable* if $\partial_\alpha f\big|_x \neq \emptyset$ for all $x \in \Omega$.

PROPOSITION 2.1 (Rolewicz, 2001) *Let* Ω *be an open convex set in a Banach space* X*. Let* $f(\cdot)$ *be a strongly* $\alpha(\cdot)$*-paraconvex function. Then its* $\alpha(\cdot)$*-subdifferential is equal to the Clarke subdifferential.*

Since by Proposition 1.2 the function $f(\cdot)$, considered in the Proposition 2.1, is locally Lipschitz and locally Lipschitz functions have non-empty Clarke subdifferentials at each point, we get:

COROLLARY 2.1 ((Rolewicz, 2001)) *Let* Ω *be an open convex set in a Banach space* X*. Let* $f(\cdot)$ *be a strongly* $\alpha(\cdot)$*-paraconvex function. Then it is* $\alpha(\cdot)$*-subdifferentiable.*

In a similar way we can consider a uniform Fréchet gradient. Namely we say that $x^* \in X^*$ is an $\alpha(\cdot)$-*gradient* of the function $f(\cdot)$ at x if

$$|f(x+h) - f(x) - x^*(h)| \leq \alpha(\|h\|). \tag{20}$$

By linearity of x^* and property (2) of $\alpha(\cdot)$ the $\alpha(\cdot)$-gradient is unique.

We say that a function $f(x)$ is $\alpha(\cdot)$-*differentiable* if it has $\alpha(\cdot)$-gradient for all $x \in \Omega$.

3. The $\alpha(\cdot)$-differentiability of strongly $\alpha(\cdot)$-paraconvex functions

At the beginning of this section we shall show:

THEOREM 3.1 (Rolewicz, 2002) *Let* Ω *be an open convex set in a separable Banach space* X*. Let* $f(\cdot)$ *be an* $\alpha(\cdot)$*-subdifferentiable function defined on* Ω*. Suppose that the dual space* X^* *is separable. Then there is a residual set* $D \subset \Omega$ *such that the function* $f(\cdot)$ *is Fréchet differentiable at every point* $x_0 \in D$*. Moreover, on* D *the Fréchet gradient is continuous in the norm in conjugate space* X^**.*

The proof is based on several notions. The first one is the notion of $\alpha(\cdot)$-monotonicity of multifunctions.

Let, as before, $\alpha(t)$ be a function mapping the interval $[0, +\infty)$ into the interval $[0, +\infty]$ such that $\alpha(0) = 0$ and such that

$$\lim_{t \downarrow 0} \frac{\alpha(t)}{t} = 0. \tag{2}$$

We say that a multifunction Γ mapping X into X^* is $\alpha(\cdot)$-*monotone* if for all $x^* \in \Gamma(x)$, $y^* \in \Gamma(y)$ we have

$$[y^* - x^*](y - x) + \alpha(\|x - y\|) \geq 0. \tag{21}$$

In particular case $\alpha(t) \equiv 0$ we obtain the classical definition of *monotone multifunctions*. The notion of $\alpha(\cdot)$-monotonicity can be considered as a uniformization of submonotonicity introduced by Springarn (1981, 1981-2).

In the case when $\alpha(t) = t^\gamma$ we obtain the definition of a γ-monotone multifunctions introduced by Jourani (1996).

It is obvious that, if a function $f(x)$ is $\alpha(\cdot)$-subdifferentiable, then its $\alpha(\cdot)$-subdifferential $\partial_\alpha f\big|_x$ considered as a multifunction of x is $2\alpha(\cdot)$-monotone.

Indeed, take arbitrary $x, y \in X$. Let $x^* \in \partial_\alpha f\big|_x$, $y^* \in \partial_\alpha f\big|_y$. By definition

$$f(y) - f(x) \geq x^*(y - x) - \alpha(\|y - x\|) \tag{22}$$

and

$$f(x) - f(y) \geq y^*(x - y) - \alpha(\|y - x\|). \tag{23}$$

Adding (22) and (23) we finally obtain

$$0 \geq [x^* - y^*](y - x) - 2\alpha(\|x - y\|). \tag{24}$$

Thus

$$0 \leq [x^* - y^*](x - y) + 2\alpha(\|x - y\|). \tag{24'}$$

Following Preiss and Zajiček (1984) we denote for any $x^* \in X^*$, $0 < \beta < 1$, $x \in X$,

$$K(x^*, \beta, x) = \{y \in X : x^*(y - x) \geq \beta\|x^*\|\|y - x\|\}. \tag{25}$$

The set $K(x^*, \beta, x)$ will be called an β*-cone with vertex at x and direction x^**.

Now we shall extend a little this definition. Namely, let $\varrho > 0$ the set

$$K(x^*, \beta, x, \varrho) = K(x^*, \beta, x) \cap \{y : \|x - y\| < \varrho\}$$

will be called an (β, ϱ)*-cone with vertex at x and direction x^**. It is obvious that the set $K(x^*, \beta, x)$ has a nonempty interior and, even more,

$$x \in \overline{\mathrm{Int} K(x^*, \beta, x, \varrho)}. \tag{26}$$

Observe that just from the definition it follows that if $\beta_1 < \beta_2$, then

$$K(x^*, \beta_1, x, \varrho) \supset K(x^*, \beta_2, x, \varrho).$$

We recall that $M \subset X$ is said to be *β-cone meagre* if for every $x \in M$ and arbitrary $\varepsilon > 0$ there are $z \in X$, $\|x - z\| < \varepsilon$ and $x^* \in X^*$ such that

$$M \cap \text{Int } K(x^*, \beta, z) = \emptyset \tag{27}$$

(Preiss and Zajiček, 1984).

Similarly, a set $M \subset X$ is said to be *(β, ϱ)-cone meagre* if for every $x \in M$ and arbitrary $\varepsilon > 0$ there are $z \in X$, $\|x - z\| < \varepsilon$ and $x^* \in X^*$ such that

$$M \cap \text{Int } K(x^*, \beta, z, \varrho) = \emptyset. \tag{28}$$

The arbitrariness of ε and (28) implies that an (β, ϱ)-cone meagre set M is nowhere dense.

We recall that a set $M \subset X$ is called *angle-small* if it can be represented as a union of a countable number of β-cone meagre sets M_n,

$$M = \bigcup_{n=1}^{\infty} M_n. \tag{29}$$

We say that $M \subset X$ is *weakly angle-small* if it can be represented as a union of a countable number of (β, ϱ_n)-cone meagre sets M_n,

$$M = \bigcup_{n=1}^{\infty} M_n \tag{30}$$

for certain $\beta > 0$.

Of course, every weakly angle-small set M is of the first Baire category.

Adapting the method of Preiss and Zajiček (1984) we obtain:

THEOREM 3.2 *Let $(X, \|\cdot\|)$ be a Banach space. Let $\Omega \subset X$ be a convex set with non-empty interior. Assume that X^* is separable. Let a multifunction Γ mapping X into 2^{X^*} be $\alpha(\cdot)$-monotone and such that* dom $\Gamma = \Omega$ *(i.e., $\Gamma(x) \neq \emptyset$ for all $x \in \Omega$). Then there exists a weakly angle-small set A such that Γ is single-valued and continuous on the set $\Omega \setminus A$.*

Proof. It is sufficient to show that the set

$$A = \{x \in X : \lim_{\delta \to 0} \text{diam } \Gamma(B(x, \delta)) > 0\}, \tag{31}$$

where diam denotes the diameter of the set, is weakly angle-small. Of course, we can represent the set A as a union of sets

$$A_n = \{x \in X : \lim_{\delta \to 0} \text{diam } \Gamma(B(x, \delta)) > \frac{1}{n}\}. \tag{32}$$

Let $\{x_m^*\}$ be a dense sequence in the space X^*. Let

$$A_{n,m} = \{x \in A_n : \text{dist}(x_m^*, \Gamma(x)) < \frac{\beta}{4n}\}, \tag{33}$$

where, as usual we denote $\mathrm{dist}(x_m^*, \Gamma(x)) = \inf\{\|x_m^* - x^*\| : x^* \in \Gamma(x)\}$. By the density of the sequence $\{x_m^*\}$ in X^*

$$\bigcup_{m=1}^{\infty} A_{n,m} = A_n.$$

We will show that the sets $A_{n,m}$ are (β, ϱ)-meagre for sufficiently small ϱ.

Indeed, suppose that $x \in A_{n,m}$. Let ε be an arbitrary positive number. Since $x \in A_n$, by (32), the are $0 < \delta < \varepsilon$ and $z_1, z_2 \in \Omega$, $x_1^* \in \Gamma(z_1)$, $x_2^* \in \Gamma(z_2)$ such that $d(z_1, x) < \delta$, $d(z_2, x) < \delta$ and

$$\|x_1^* - x_2^*\| > \frac{1}{n}. \tag{34}$$

Thus by the triangle inequality, for every $x^* \in \Gamma(x)$ either $\|x_1^* - x^*\| > \frac{1}{2n}$ or $\|x_2^* - x^*\| > \frac{1}{2n}$. By the definition of $A_{n,m}$, we can find $x_x^* \in \Gamma(x)$ such that $\|x_x^* - x_m^*\| < \frac{\beta}{4n}$. Therefore choosing as z either z_1 or z_2, we can say that there are points $z \in X$ and $x_z^* \in \Gamma(z)$ such that $d(z, x) < \delta$ and

$$\|x_z^* - x_m^*\| \geq \|x_z^* - x_x^*\| - \|x_x^* - x_m^*\| > \frac{1}{2n} - \frac{\beta}{4n} > \frac{1}{4n}. \tag{35}$$

Since (2) there is ϱ_n such that

$$\frac{1}{2n} - \frac{\beta}{4n} - \frac{1}{\beta} \sup_{0<t<\varrho_n} \frac{\alpha(t)}{t} > \frac{1}{4n}. \tag{36}$$

We shall show that

$$A_{n,m} \cap K(x_z^* - x_m^*, \beta, z) \cap \{y : \|z - y\| < \varrho_n\} = \emptyset.$$

Indeed, suppose that $y \in K(x_z^* - x_m^*, \beta, z)$. This means that

$$[x_z^* - x_m^*](y - z) \geq \beta \|x_z^* - x_m^*\| \|y - z\|.$$

Suppose that $x_y^* \in \Gamma(y)$. Since Γ is $\alpha(\cdot)$-monotone by definition we have

$$[x_y^* - x_z^*](y - z) \geq -\alpha(\|y - z\|).$$

Adding this two inequalities we get

$$[x_y^* - x_m^*](y - z) \geq \beta \|x_z^* - x_m^*\| \|y - z\| - \alpha(\|y - z\|).$$

and if additionally $\|y - z\| < \varrho_n$ we have by (36)

$$\begin{aligned}
&[x_y^* - x_m^*](y - z) \geq \beta \|x_z^* - x_m^*\| \|y - z\| - \alpha(\|y - z\|) \\
&\geq \beta\Big(\frac{1}{2n} - \frac{\beta}{4n}\Big)\|y - z\| - \alpha(\|y - z\|) \geq \beta\Big(\frac{1}{2n} - \frac{\beta}{4n} - \frac{1}{\beta}\frac{\alpha(\|y - z\|)}{\|y - z\|}\Big)\|y - z\| \\
&> \frac{\beta}{4n}\|y - z\|.
\end{aligned}$$

This implies that

$$\|x_y^* - x_m^*\| > \frac{\beta}{4n}$$

and by the definition of $A_{n,m}$, $y \notin A_{n,m}$. Thus

$$A_{n,m} \cap K(x_z^* - x_m^*, \beta, z) \cap \{y : \|z - y\| < \varrho_n\} = \emptyset$$

and the set $A_{n,m}$ is (β, ϱ_n)-meagre. Therefore the set A is weakly angle-small. ■

Since the subdifferential $\partial_\alpha f\big|_x$ of an $\alpha(\cdot)$-subdifferentiable function is a $2\alpha(\cdot)$-monotone multifunction of x, we immediately obtain:

COROLLARY 3.1 *Let $(X, \|\cdot\|)$ be a Banach space. Let $\Omega \subset X$ be a convex set with non-empty interior. Assume that X^* is separable. Let $f(x)$ be an $\alpha(\cdot)$-subdifferentiable function defined on Ω. Then there is a weakly angle-small set A such that on the set $\Omega \setminus A$ the $\alpha(\cdot)$-subdifferential $\partial_\alpha f\big|_x$ is single-valued and continuous.*

Since the weakly angle-small sets are always of the first Baire category we immediately obtain:

COROLLARY 3.2 *Let $(X, \|\cdot\|)$ be a Banach space. Let $\Omega \subset X$ be a convex set with non-empty interior. Assume that X^* is separable. Let $f(x)$ be an $\alpha(\cdot)$-subdifferentiable function defined on Ω. Then there is a residual set $D \subset \Omega$ such that on the set of D the $\alpha(\cdot)$-subdifferential $\partial_\alpha f\big|_x$ is single-valued and continuous.*

Proof of Theorem 3.1. Recall that, if at a point x_0 the Clarke subgradient $\ell = \partial^{Cl} f|_{x_0}$ is unique, than this gradient is a Gateaux differential. Indeed, let h be fixed. By simplicity we denote $\tilde{f}(t) = f(x_0 + th) - f(x_0) - t\ell(h)$. The fact that at a point x_0 the Clarke subgradient $\ell = \partial^{Cl} f|_{x_0}$ is unique is nothing else than the fact that

$$\limsup_{\substack{t \to 0 \\ x \to x_0}} \frac{\tilde{f}(x + th) - \tilde{f}(x)}{t} = 0. \tag{37}$$

Suppose that the function $\tilde{f}(\cdot)$ is not differentiable at 0. It means that there are $a > 0$ and a sequence $\{t_n\}$ tending to 0 such that

$$\tilde{f}(t_n) \leq -a|t_n|. \tag{38}$$

Replacing eventually $\tilde{f}(t)$ by $\tilde{f}(-t)$ we can assume without loss of generality, that $t_n > 0$. Let $x_n = t_n h$. Obviously $0 = x_n - t_n h$ and by (38)

$$\liminf_{\substack{t_n \to 0 \\ x_n \to 0}} \frac{\tilde{f}(0) - \tilde{f}(x_n)}{t_n} \geq a, \tag{39}$$

which contradicts (37).

Thus by the classical theorem that a continuous Gateaux differential is a Fréchet differential from Corollaries 3.1 and 3.2 we trivially obtain Theorem 3.1. ■

Now we shall show how to extend Theorem 3.1 to the case when X is a non-separable Asplund space. The proof (Rolewicz, 2005) is similar as the proof that X is an Asplund space if and only if each of its separable subspaces is an Asplund space.

PROPOSITION 3.1 *Let* $(X, \|\cdot\|)$ *be a real Banach space. Let* $f(\cdot)$ *be a function defined on an open convex subset* $\Omega \subset X$. *Suppose that* x^* *is an approximate subgradient of the function* $f(\cdot)$ *in* $x \in \Omega$. *Then* x^* *is the Fréchet gradient of the function* $f(\cdot)$ *at the point* x *if and only if for arbitrary* $\varepsilon > 0$ *there is* $\delta > 0$ *such that*

$$\frac{f(x+ty)+f(x-ty)-2f(x)}{t} < \varepsilon \tag{40}$$

for all $y \in X$ *such that* $\|y\| = 1$ *and* $0 < t < \delta$.

If $f(\cdot)$ is strongly $\alpha(\cdot)$-paraconvex we can replace the request that (40) holds by t small enough by the fact that such t exists and we obtain:

PROPOSITION 3.2 *Let* $(X, \|\cdot\|)$ *be a real Banach space. Let* $f(\cdot)$ *be a strongly* $\alpha(\cdot)$*-paraconvex function defined on an open convex subset* $\Omega \subset X$. *Then the function* $f(\cdot)$ *is Fréchet differentiable at a point* $x \in \Omega$ *if and only if for arbitrary* $\varepsilon > 0$ *there is* $t_\varepsilon > 0$ *such that*

$$\frac{f(x+t_\varepsilon y)+f(x-t_\varepsilon y)-2f(x)}{t_\varepsilon} < \varepsilon \tag{41}$$

for all $y \in X$ *such that* $\|y\| = 1$.

As a consequence we get that the set G (possibly empty) of points $x \in \Omega$ where the function $f(\cdot)$ is Fréchet differentiable is a G_δ set. Therefore, if the set G of points $x \in \Omega$ where the function $f(\cdot)$ is Fréchet differentiable is dense in Ω, then it is a residual set.

Now, suppose that $f(\cdot)$ is a strongly $\alpha(\cdot)$-paraconvex function defined on an open convex subset $\Omega \subset X$ and that the set G of points $x \in \Omega$ where the function $f(\cdot)$ is Fréchet differentiable is **not** dense in Ω. Using the construction given in the proof that X is an Asplund space if and only if each its separable subspaces is an Asplund space, we can show that there is a separable subspace $E \subset X$ such that the points of Fréchet differentiability of the restriction of the function $f(\cdot)$ to $\Omega \cap E$, $f|_{\Omega\cap E}$, is **not** dense in $\Omega \cap E$.

This fact together with Theorem 3.1 gives Theorem 1.2.

4. Extensions to metric spaces

The results of Section 3 can be extended to metric spaces. In fact, I ought to say that at the beginning the results were formulated in this more general setting.

Let X be a metric space. Let Φ be an arbitrary family of functions defined on X and having values in the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. Let $f(\cdot)$ be a real-valued function defined on X. Similarly as in the classical case, a function $\phi(\cdot) \in \Phi$ will be called a Φ-*subgradient* of the function $f(\cdot)$ at a point x_0 if

$$f(x) - f(x_0) \geq \phi(x) - \phi(x_0) \tag{42}$$

for all $x \in X$. The set of all Φ-subgradients of the function $f(\cdot)$ at a point x_0 shall be called the Φ-*subdifferential* of the function f at the point x_0 and we shall denote it by $\partial_\Phi f\big|_{x_0}$.

Of course $\partial_\Phi f\big|_{(\cdot)}$ is a multifunction mapping X into 2^Φ. It is not too difficult to observe that this multifunction is cyclic monotone, i.e. for arbitrary n and $x_0, x_1, ..., x_n = x_0 \in X$ and $\phi_{x_i} \in \partial_\Phi f\big|_{x_i}$, $i = 0, 1, 2, ..., n$, we have

$$\sum_{i=1}^{n} [\phi_{x_{i-1}}(x_{i-1}) - \phi_{x_{i-1}}(x_i)] \geq 0. \tag{43}$$

We shall say that a function $f(\cdot)$ mapping a metric space (X, d_X) into $\mathbb{R}$ is *Fréchet* Φ-*differentiable* at a point x_0 if there is a function $\phi \in \Phi$ such that

$$\lim_{x \to x_0} \frac{|[f(x) - f(x_0)] - [\phi(x) - \phi(x_0)]|}{d_X(x, x_0)} = 0. \tag{44}$$

The function ϕ will be called a *Fréchet* Φ-*gradient* of the function $f(\cdot)$ at the point x_0. The set of all Fréchet Φ-gradients of the function $f(\cdot)$ at the point x_0 is called *Fréchet* Φ-*differential* of the function $f(\cdot)$ at the point x_0 and it is denoted by $\partial_\Phi^F f\big|_{x_0}$.

Under proper assumptions we can obtain an extension of the famous Asplund theorem to the case of metric spaces. The assumptions are as follow:

(a) Φ is an additive group,

(sL) Φ is a set of Lipschitz functions. Moreover the space $\Phi/\mathbb{R}$ is separable in the Lipschitz norm $\|\phi\|_L$,

(wm) the family Φ has the *weak k-monotonicity property*, $0 < k \leq 1$, i.e. for all $x \in X$, $\phi \in \Phi$ and $t > 0$, there is a $y \in X$ such that $0 < d_X(x, y) < t$ and

$$|\phi(y) - \phi(x)| \geq k\|\phi\|_L d_X(y, x). \tag{45}$$

THEOREM 4.1 (Rolewicz, 2002) *Let X be a metric space. Let Φ be a family of Lipschitz functions satisfying assumptions* (a), (sL) *and* (wm). *Let a multifunction Γ mapping X into 2^Φ be monotone and such that* dom $\Gamma = X$ (*i.e.,*

$\Gamma(x) \neq \emptyset$ for all $x \in X$). Then there exists a residual set Ω such that Γ is single-valued and continuous (i.e. simultaneously lower semi-continuous and upper semi-continuous) at each point of Ω.

Recall that in the case of normed spaces Gateaux differentiability of a convex continuous functions $f(\cdot)$ at a point x is equivalent to the fact that the subdifferential $\partial f\big|_x$ consists of one point only. Moreover the continuity of Gateaux differentials in the norm operator topology implies that these differentials are the Fréchet differentials. Similarly we have an extension of this fact to metric spaces (Rolewicz, 1995, 1996). As a consequence we get:

THEOREM 4.2 (Rolewicz, 2002) *Let X be a metric space, which is of the second Baire category on itself (in particular, let X be a complete metric space). Let Φ be a family of Lipschitz functions satisfying assumptions* (a), (sL) *and* (wm). *Let $f(\cdot)$ be a continuous Φ-subdifferentiable function. Then there is a residual set Ω such that the function $f(\cdot)$ is Fréchet Φ-differentiable at every point $x_0 \in \Omega$. Moreover, on Ω the Fréchet Φ-gradient is unique and it is continuous in the metric d_L.*

References

ASPLUND, E. (1966) Farthest points in reflexive locally uniformly rotund Banach spaces. *Israel Jour. Math.* **4**, 213-216.

ASPLUND, E. (1968) Fréchet differentiability of convex functions. *Acta Math.* **121**, 31-47.

FABIAN, M. (1989) Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. *Acta Univ. Carolinae* **30**, 51-56.

IOFFE, A.D. (1984) Approximate subdifferentials and applications I. *Trans. AMS* **281**, 389-416.

IOFFE, A.D. (1986) Approximate subdifferentials and applications II. *Mathematika* **33**, 111-128.

IOFFE, A.D. (1989) Approximate subdifferentials and applications III. *Mathematika* **36**, 1-38.

IOFFE, A.D. (1990) Proximal analysis and approximate subdifferentials. *J. London Math. Soc.* **41**, 175-192.

IOFFE, A.D. (2000) Metric regularity and subdifferential calculus, in Russian. *Usp. Matem. Nauk* **55** (3), 104-162.

JOFRÉ A., LUC D.T. and THÉRA, M. (1998) ε-subdifferential and ε-monotonicity. *Nonlinear Analysis, Theory, Methods and Appl.* **33**, 71-90.

JOURANI, A. (1996) Subdifferentiability and subdifferential monotonicity of γ-paraconvex functions. *Control and Cyber.* **25**, 721-737.

Luc D.T., Ngai H.V. and Théra, M. (1999) On ε-convexity and ε-monotonicity. In: A.Ioffe, S.Reich and I. Shapiro, eds., *Calculus of Variation and Differential Equations, Research Notes in Mathematics Series*, Chapman and Hall, 82-100.

Luc D.T., Ngai H.V. and Théra, M. (2000) Approximate convex functions. *J. Nonlinear and Convex Anal.* **1**, 155-176.

Mazur, S. (1933) Über konvexe Mengen in linearen normierten Räumen. *Stud. Math.* **4**, 70-84.

Mordukhovich, B.S. (1976) Maximum principle in the optimal control problems with non-smoth constraints, in Russian. *Prikl. Mat. Meh.* **40**, 1014-1023.

Mordukhovich, B.S. (1980) Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, in Russian. *Soviet Math. Doklady* **254**, 1072-1076. In English version **22**, 526-530.

Mordukhovich, B.S. (1988) *Approximation Methods in Problems of Optimization and Control* (in Russian). Nauka, Moscow.

Pallaschke, D. and Rolewicz, S. (1997) *Foundation of mathematical optimization.* Mathematics and its Applications 388, Kluwer Academic Publishers, Dordrecht/Boston/London, 1997.

Phelps, R.R. (1989) *Convex Functions, Monotone Operators and Differentiability.* Lecture Notes in Mathematics, Springer-Verlag **1364**.

Preiss, D. and Zajíček, L. (1984) Stronger estimates of smallness of sets of Fréchet nondifferentiability of convex functions. *Proc. 11-th Winter School, Suppl. Rend. Circ. Mat di Palermo, ser II*, **3**, 219-223.

Rockafellar, R.T. (1970) *Convex Analysis.* Princeton University Press.

Rockafellar, R.T. (1980) Generalized directional derivatives and subgradient of nonconvex functions. *Can. Jour. Math.* **32**, 257-280.

Rockafellar, R.T. (1982) Favorable classes of Lipschitz continuous functions in subgradient optimization. In: E. Nurminski, ed., *Nondifferentiable optimization*, Pergamon Press, New York.

Rolewicz, S. (1979a) On paraconvex multifunctions. *Oper. Research Verf.* (Methods of Oper. Res.) **31**, 540-546.

Rolewicz, S. (1979b) On γ-paraconvex multifunctions. *Math. Japonica* **24**, 293-300.

Rolewicz, S. (1980) On conditions warranting Φ_2-subdifferentiability. *Studies in Math. Programming* **14**, 215-224.

Rolewicz, S. (1993) On Asplund inequalities for Lipschitz functions. *Arch. der Math.* **61**, 484-488.

Rolewicz, S. (1994) On Mazur Theorem for Lipschitz functions. *Arch. der Math.* **63**, 535-540.

Rolewicz, S. (1995a) On Φ-differentiability of functions over metric spaces. *Topological Methods of Nonlinear Analysis* **5**, 229-236.

Rolewicz, S. (1995b) On subdifferential on non-convex sets. In: D. Przeworska-Rolewicz, ed., *Different Aspects of Differentiability, Dissertationes Math.* **340**, 301-308.

Rolewicz, S. (1999) On $\alpha(\cdot)$-monotone multifunction and differentiability of γ-paraconvex functions. *Stud. Math.* **133**, 29-37.

Rolewicz, S. (2000) On $\alpha(\cdot)$-paraconvex and strongly $\alpha(\cdot)$-paraconvex functions. *Control and Cybernetics* **29**, 367-377.

Rolewicz, S. (2001a) On the coincidence of some subdifferentials in the class of $\alpha(\cdot)$-paraconvex functions. *Optimization* **50**, 353-360.

Rolewicz, S. (2001b) On uniformly approximate convex and strongly $\alpha(\cdot)$-paraconvex functions. *Control and Cybernetics* **30**, 323-330.

Rolewicz, S. (2002) $\alpha(\cdot)$-monotone multifunctions and differentiability of strongly $\alpha(\cdot)$-paraconvex functions. *Control and Cybernetics* **31**, 601-619.

Rolewicz, S. (2005) On differentiability of strongly $\alpha(\cdot)$-paraconvex functions in non-separable Asplund spaces. *Studia Math.* **167**, 235-244.

Spingarn, J.E. (1981) Submonotone subdifferentials of Lipschitz functions. *Trans. Amer. Math. Soc.* **264**, 77-89.

Spingarn, J.E. (1981-2) Submonotone mappings and the proximal point algorithm. *Numer. Funct. Anal. Opt.* **4** 123-150.