Topological **M**ethods in **N**onlinear **A**nalysis
Journal of the Juliusz Schauder Center
Volume 5, 1995, 229–236

ON Φ-DIFFERENTIABILITY OF FUNCTIONS OVER METRIC SPACES

S. Rolewicz

Dedicated to Ky Fan

In 1933 S. Mazur [4] proved the following

Theorem 1. *Let $(X, \|\cdot\|)$ be a separable real Banach space. Let f be a real-valued convex continuous function defined on an open convex subset $\Omega \subset X$. Then there is a subset $A \subset \Omega$ of the first category such that f is Gateaux differentiable on $\Omega \setminus A$.*

The result of Mazur was a starting point for the theory of differentiability of convex functions (cf. the book of Phelps [7]).

In 1968 Asplund showed that under the assumption that in the conjugate space X^* there is an equivalent strictly convex norm (in particular, if X^* is separable) we can obtain the Fréchet differentiability.

Theorem 2 ([1]). *Let $(X, \|\cdot\|)$ be a real Banach space. Suppose that in the conjugate space X^* there is an equivalent strictly convex norm. Let f be a real-valued convex continuous function defined on an open convex subset $\Omega \subset X$. Then there is a subset $A \subset \Omega$ of the first category such that f is Fréchet differentiable on $\Omega \setminus A$.*

A question arises how to extend this theorem to functions defined on a metric space without any linear structure.

1991 *Mathematics Subject Classification.* Primary 47H05, 47N10; Secondary 47B47, 54E35.
Key words and phrases. Metric spaces, Φ-subdifferentiability, monotone operators.

The first step is, of course, to extend the notion of convexity. The idea of extending this notion to nonlinear spaces is fairly old. The first step in this direction was made by Menger [6]. Menger considered a metric space (X,d). He called a set $A \subset X$ *convex* if for every $x, y \in A$, $0 \le t \le 1$, and every z such that $d(x,z) = td(x,y)$ and $d(y,z) = (1-t)d(x,y)$, we have $z \in A$. Using this metric approach to convexity, Mazur and Ulam [5] proved that every isometric mapping $T : X \to Y$ of a real Banach space X onto a real Banach space Y such that $T(0) = 0$ is a linear mapping.

In this paper, however, we shall concentrate on another approach to extension of convexity. This approach started with the paper of Ellis [2], who investigated the separation property by functions which are not linear. There is a very nice result of Ky Fan [3], who extended the Krein–Milman theorem to the so-called Φ-convex compact sets. Later this direction was developed in a series of papers of different authors under the name of *axiomatic convexity*. The reader can find the description of that theory in the book of Soltan [12].

Now we describe the Ellis–Ky Fan approach. Let an arbitrary set X, called later *the space*, be given. Let Φ be a collection of functions defined on X with values in $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

A function $\phi \in \Phi$ will be called a Φ-*subgradient* of the function $f : X \to \overline{\mathbb{R}}$ at a point x_0 if

$$f(x) - f(x_0) \ge \phi(x) - \phi(x_0). \tag{1}$$

The set of all Φ-subgradients of a function f at a point x_0 will be called the Φ-*subdifferential* of f at x_0 and denoted by $\partial_\Phi f|_{x_0}$.

Observe that the order in the real numbers induces a partial order in the real-valued functions in the following way. We write $g \le f$, omitting the argument, if $g(x) \le f(x)$ for all $x \in X$. For a given function f we set

$$f^\Phi(x) = \sup\{\phi(x) + c : \phi \in \Phi,\ c \in \mathbb{R},\ \phi + c \le f\}. \tag{2}$$

The function f^Φ is called the Φ-*convexification* of f. If $f^\Phi = f$ we say that f is Φ-*convex*. We denote the set of all Φ-convex functions by Φ_{conv}.

In the sequel we assume that X is a metric space with metric d_X. The first step is to define Fréchet and Gateaux differentiable functions in this case.

Fréchet differentiability is easily defined. A function $\phi \in \Phi$ is called a *Fréchet* Φ-*gradient* of a real-valued function f at a point x_0 if for every $\varepsilon > 0$, there is a neighbourhood U of x_0 such that for $x \in U$,

$$|[f(x) - f(x_0)] - [\phi(x) - \phi(x_0)]| < \varepsilon d(x, x_0). \tag{3}$$

We say that a real-valued function defined on a metric space is *Fréchet* Φ-*differentiable* at x_0 if there exists a Fréchet Φ-gradient of f at x_0. In general, there may be more than one such Fréchet Φ-gradient. The set of all Fréchet

Φ-gradients of f at x_0 will be called the *Fréchet Φ-differential* of f at x_0 and denoted by $\partial_\Phi^{\mathrm{F}} f|_{x_0}$.

The definition of Fréchet Φ-gradients and Fréchet Φ-differentials of a function f at a point x_0 can be extended to functions mapping a metric space (X, d_X) into a metric space (Y, d_Y) in the following way. Let metric spaces (X, d_X) and (Y, d_Y) be given. Let Φ be a collection of mappings of X into Y. A function $\phi \in \Phi$ is called a *Fréchet Φ-gradient* of a function $f : X \to Y$ at a point $x_0 \in X$ if $f(x_0) = \phi(x_0)$ and for every $\varepsilon > 0$, there is a neighbourhood U of x_0 such that for $x \in U$,

$$d_Y(f(x), \phi(x)) < \varepsilon d(x, x_0). \tag{3$_m$}$$

We say that $f : X \to Y$ is *Fréchet Φ-differentiable* at x_0 if there exists a Fréchet Φ-gradient of f at x_0.

In this paper, however, we restrict ourselves to real-valued functions.

In the case of Gateaux differentiable functions we do not have a good definition in the general case, even for real-valued functions. However, in the case of Φ-convex functions such a definition can be given.

Observe that a convex function of one variable is differentiable if and only if the subgradient at x_0 is unique. As a consequence, a continuous convex function f defined on a linear space is Gateaux differentiable at x_0 if and only if the subgradient of f at x_0 is uniquely determined.

Having this in mind, we can reformulate the Mazur theorem [4] in the following way:

Theorem 1′. *Let $(X, \|\cdot\|)$ be a separable real Banach space. Let f be a real-valued convex continuous function defined on an open convex subset $\Omega \subset X$. Then there is a subset $A \subset \Omega$ of the first category such that the subgradient of f is uniquely determined on $\Omega \setminus A$.*

In order to extend Theorem 1′ to the nonlinear case, we need the notion of monotone multifunctions. A multifunction Γ mapping X into 2^Φ is said to be *monotone* if for any $\phi_x \in \Gamma(x)$ and $\phi_y \in \Gamma(y)$ we have

$$\phi_x(x) + \phi_y(y) - \phi_x(y) - \phi_y(x) \geq 0. \tag{4}$$

In the particular case when X is a linear space, and Φ is a linear space consisting of linear functionals $\phi(x)$, denoted by $\phi(x) = \langle \phi, x \rangle$, we can rewrite (4) in the classical form

$$\langle \phi_x - \phi_y, x - y \rangle \geq 0. \tag{5}$$

As a trivial consequence of this definition, we see that for a given function f the subdifferential $\partial_\Phi f|_x$ is a monotone multifunction of x.

It is interesting to find conditions on the metric space (X, d) and on the class Φ of real-valued functions to guarantee that for any monotone multifunction

$\Gamma : X \to 2^{\Phi}$, there is a set $A_{\Gamma} \subset X$ of the first category such that outside A_{Γ} the multifunction Γ is single-valued. For this purpose we introduce some new notions.

Let (X,d) be a metric space. Let Φ be a subclass of the space of all Lipschitz functions defined on X. Let

$$d_{\mathrm{L}}(\phi_1,\phi_2) = \sup_{x_1,x_2\in X,\, x_1\neq x_2} \frac{|[\phi_1(x_1)-\phi_2(x_1)]-[\phi_1(x_2)-\phi_2(x_2)]|}{d_X(x_1,x_2)}. \tag{6}$$

It is easy to see that d_{L} is a quasimetric, i.e. it is symmetric and satisfies the triangle inequality. Observe that $d_{\mathrm{L}}(\phi_1,\phi_2)=0$ implies that the difference of ϕ_1 and ϕ_2 is a constant function, i.e. $\phi_1(x)=\phi_2(x)+c$. Thus d_{L} is a metric on the quotient space $\Phi/\mathbb{R}$.

A family Φ of locally Lipschitz functions is called *locally separable* if for every $x\in X$ there is a neighbourhood U of x such that the family $\Phi|_U$ of restrictions to U of functions from Φ is separable in the Lipschitz metric.

Let Φ be a linear family of Lipschitz functions. If there is a constant k, $0<k<1$, such that for all $x\in X$ and all $\phi\in\Phi$ and all $t>0$ there is a $y\in X$ such that $0<d_X(x,y)<t$ and

$$\phi(y)-\phi(x)\geq k d_{\mathrm{L}}(\phi,0)d_X(y,x), \tag{7}$$

we say that the family Φ has the *k-monotonicity property.* It is obvious that the restrictions to an open set U of linear continuous functionals over a real Banach space have the k-monotonicity property with any constant k, $0<k<1$. We say that Φ has the *local monotonicity property* if for every $x\in X$ there is a neighbourhood U of x such that the family $\Phi|_U$ has the k_U-monotonicity property for some k_U, $0<k_U<1$.

By an extension of the method of Preiss and Zajíček [8] to metric spaces and repeating step by step the considerations from [9], we can obtain a local version of the results of [9]:

Theorem 3. *Let X be an open set in a complete separable connected metric space. Let Φ be a linear family of locally Lipschitz functions having the local monotonicity property. Assume that Φ is locally separable. Let Γ be a monotone multifunction mapping X into Φ such that $\Gamma(x)\neq\emptyset$ for all $x\in X$. Then there exists a set $A\subset X$ of the first category such that Γ is single-valued and continuous on $X\setminus A$.*

Since the subdifferential $\partial_{\Phi}f|_x$ is a monotone multifunction of x, we immediately obtain

Corollary 4. *Let X be an open set in a complete metric space. Let Φ be a linear family of locally Lipschitz functions having the local monotonicity property. Assume that Φ is locally separable. Let f be a Φ-convex function having a Φ-subgradient at each point. Then there is a set $A\subset X$ of the first category such* *that outside A the subdifferential $\partial_\Phi f|_x$ is single-valued and continuous in the metric d_{L}.*

In the case when X is a Banach space and $\Phi = X^*$ is the space of all linear continuous functionals, Φ has the k-monotonicity property with any $k < 1$. Thus we have

Corollary 5. *Let X be a Banach space with separable dual X^*. Let f be a locally convex continuous function defined on an open set U. Then there is a set A of the first category such that on $U \setminus A$ the subdifferential $\partial f|_x$ is single-valued and continuous in the norm topology.*

Remark 6. Observe that in the case when the set U is not convex, there are locally convex functions which are not restrictions of convex functions (cf. for instance [10], [11]).

In order to obtain a result similar to the Asplund theorem, recall the following classical statement: the continuity of the Gateaux differential $dF(x,h)$ of a function $F(x)$ with respect to x in the norm operator topology implies that $dF(x,h)$ is the Fréchet differential of $F(x)$. Its metric analogue is the following:

Proposition 7. *Let (X, d_X) be a metric space. Let Φ denote a linear class of Lipschitz functions defined on X. Let f be a Φ-convex function on X. Let ϕ_{x_0} be a Φ-subgradient of f at a point $x_0 \in X$. Suppose that there is a neighbourhood U of x_0 such that for all $x \in U$, the subdifferential $\partial_\Phi f|_x$ is not empty and lower semicontinuous at x_0 in the Lipschitz norm, i.e. for every $\varepsilon > 0$ there is a neighbourhood $V_\varepsilon \subset U$ such that for $x \in V_\varepsilon$ there is $\phi_x \in \partial_\Phi f|_x$ such that*

$$\|\phi_x - \phi_{x_0}\|_L \le \varepsilon d_X(x, x_0). \tag{8}$$

Then ϕ_{x_0} is a Fréchet Φ-gradient of f at x_0.

Proof. Let

$$F(x) = [f(x) - f(x_0)] - [\phi_{x_0}(x) - \phi_{x_0}(x_0)].$$

It is easy to see that $F(x_0) = 0$. Since ϕ_{x_0} is a Φ-subgradient of f at x_0, the function $F(x)$ is nonnegative. Let ε be an arbitrary positive number and let $V_\varepsilon \subset U$ be a neighbourhood of x_0 such that (8) holds for $x \in V_\varepsilon$. Since ϕ_x is a Φ-subgradient of f at $x \in V_\varepsilon$, $\psi_x = \phi_x - \phi_{x_0}$ is a Φ-subgradient of F at x. Thus

$$F(y) - F(x) \ge \psi_x(y) - \psi_x(x).$$

In particular, if $y = x_0$ then we obtain

$$F(x_0) - F(x) \ge \psi_x(x_0) - \psi_x(x). \tag{9}$$

Taking into account (8), we obtain for $x \in V_\varepsilon$,

$$0 \le F(x) \le \varepsilon d_X(x, x_0). \tag{10}$$

The arbitrariness of ε implies that 0 is a Fréchet Φ-gradient of F at x_0. Thus ϕ_{x_0} is a Fréchet Φ-gradient of f at x_0. □

If we assume that the function f is continuous, we do not need to assume that there is a neighbourhood U of x_0 such that for all $x \in U$, the subdifferential $\partial_\Phi f|_x$ is not empty. It is sufficient to assume that the subdifferential $\partial_\Phi f|_x$ is not empty on a dense set.

PROPOSITION 8. *Let (X, d_X) be a metric space. Let Φ be a linear class of Lipschitz functions defined on X. Let f be a continuous Φ-convex function on X. Let ϕ_{x_0} be a Φ-subgradient of f at a point $x_0 \in X$. Suppose that there is a dense set A in a neighbourhood U of x_0 such that for all $x \in A$, the Φ-subdifferential $\partial_\Phi f|_x$ is not empty and lower semicontinuous at x_0 in the Lipschitz norm. Then ϕ_{x_0} is a Fréchet Φ-gradient of f at x_0.*

PROOF. The proof goes along the same lines as the proof of Proposition 7. We find that $0 \le F(x) \le \varepsilon d_X(x, x_0)$ for $x \in A \cap V_\varepsilon$. Thus, by continuity of F and density of A, the same inequality holds for all $x \in U$. The rest of the proof is the same. □

Combining Proposition 7 with Theorem 3, we obtain a generalization of the Asplund theorem:

THEOREM 9. *Let X be an open set in a a complete metric space. Denote by Φ a linear locally separable space of Lipschitz functions defined on X with the local monotonicity property. Let f be a continuous Φ-convex function on X. Then there is a dense G_δ-set $A \subset X$ such that for every $x \in A$ the function f has a Fréchet Φ-differential at x consisting of one function $\phi_x \in \Phi$.*

There is a natural question about the possible extension of Proposition 7 to non-Φ-convex functions. For this purpose, we introduce Φ_ε-subgradients and Φ_{0+}-subgradients.

Let (X, d_X) be a metric space. Let Φ be a class of real-valued functions defined on X. Let f be a real-valued function on X. We say that a function $\phi \in \Phi$ is a Φ_ε-*subgradient* of f at a point $x_0 \in X$ if for all $x \in X$,

$$f(x) - f(x_0) \ge \phi(x) - \phi(x_0) - \varepsilon d_X(x, x_0). \tag{11}$$

We say that $\phi \in \Phi$ is a Φ_{0+}-*subgradient* of f at a point x_0 if for every $\varepsilon > 0$ there is a neighbourhood V_ε of x_0 such that for all $x \in V_\varepsilon$,

$$f(x) - f(x_0) \ge \phi(x) - \phi(x_0) - \varepsilon d_X(x, x_0). \tag{12}$$

The set of all Φ_{0+}-subgradients of f at x_0 is called the Φ_{0+}-*subdifferential* of f at x_0 and denoted by $\partial_{\Phi_{0+}} f|_{x_0}$. Having this notion, we can extend Proposition 7 to metric spaces under some additional assumptions.

Let a metric space (X, d_X) be given. By an *arc* in X we shall understand a homeomorphic image L of the interval $[0, 1]$, i.e. a function x defined on the interval $[0, 1]$ with values in X such that $x(t) = x(t')$ implies $t = t'$.

The points $x(0)$, $x(1)$ are called the *beginning* and the *end* of the arc, respectively. By the *length* of an arc $L = \{x(t)\}$, $0 \le t \le 1$, we mean $l(L) = \sup\{\sum_{i=1}^n d_X(x(t_i), x(t_{i-1})) : 0 = t_0 < t_1 < \ldots < t_n = 1\}$.

We say that a metric space (X, d_X) is *arc connected with constant* $K \ge 1$ if for any $x_0, x_1 \in X$ there is an arc $L = \{x(t)\}$, $0 \le t \le 1$, in X such that $x(0) = x_0$, $x(1) = x_1$ and $l(L) \le K d_X(x_0, x_1)$.

Proposition 10. *Let (X, d_X) be a metric space which is arc connected with a constant $K \ge 1$. Denote by Φ a linear class of Lipschitz functions defined on X. Let f be a real-valued function on X and let ϕ_{x_0} be a Φ_{0+}-subgradient of f at a point x_0. Suppose that there is a neighbourhood U of x_0 such that for all $x \in U$, the Φ_{0+}-subdifferential $\partial_{\Phi_{0+}} f|_x$ is nonempty and lower semicontinuous at x_0 in the Lipschitz norm, i.e. for every $\varepsilon > 0$ there is a neighbourhood $V_\varepsilon \subset U$ such that for $x \in V_\varepsilon$ there is $\phi_x \in \partial_{\Phi_{0+}} f|_x$ with the property*

$$d_{\mathrm{L}}(\phi_x, \phi_{x_0}) \le \varepsilon d_X(x, x_0). \tag{13}$$

Then ϕ_{x_0} is a Fréchet Φ-gradient of f at x_0.

Proof. Let

$$F(x) = [f(x) - f(x_0)] - [\phi_{x_0}(x) - \phi_{x_0}(x_0)].$$

It is easy to see that $F(x_0) = 0$. Since ϕ_{x_0} is a Φ_{0+}-subgradient of f at x_0, 0 is a Φ_{0+}-subgradient of F at x_0. Let $\varepsilon > 0$. Then there is a neighbourhood U_ε of x_0 such that for $x \in U_\varepsilon$,

$$F(x) \ge -\, \varepsilon d_X(x, x_0). \tag{14}$$

Let $V_\varepsilon \subset U_\varepsilon$ be a neighbourhood of x_0 such that (13) holds for $x \in V_\varepsilon$. Since ϕ_x is a Φ_{0+}-subgradient of f at x, $\psi_x = \phi_x - \phi_{x_0}$ is a Φ_{0+}-subgradient of F at x. Thus there is a neighbourhood U_ε^x of x_0 such that for $y \in U_\varepsilon^x$,

$$F(y) - F(x) \ge \psi_x(y) - \psi_x(x) - \varepsilon d_X(x, y). \tag{15}$$

Since the space (X, d_X) is arc connected with constant $K > 0$, there is a curve $L = \{x(t)\}$, $0 \le t \le 1$, with beginning x_0 and end x and with length not greater than $K d_X(x, x_0)$.

Using a compactness argument, we conclude that there are a finite number of points $x(0) = x_0$, $x(t_1), \ldots, x(t_n) = x$, $0 = t_0 < t_1 < \ldots < t_n = 1$, such that $x(t_{i-1}) \in U_\varepsilon^{x(t_i)}$. Thus by (12), for $i = 1, 2, \ldots, n$ we obtain

$$\begin{aligned} F(x(t_{i-1})) - F(x(t_i)) \ge \psi_{x(t_i)}(x(t_{i-1})) - \psi_{x(t_i)}(x(t_i)) \\ -\, \varepsilon d_X(x(t_i), x(t_{i-1})). \end{aligned} \tag{16_i}$$

Adding all equations (16_i) and taking into account that the length of L is not greater than $K d_X(x, x_0)$, we obtain

$$-\, \varepsilon d_X(x, x_0) \le F(x) \le \varepsilon 2K d_X(x, x_0). \tag{17}$$

The arbitrariness of ε implies that 0 is a Fréchet Φ-gradient of F at x_0. This trivially implies that ϕ_{x_0} is a Fréchet Φ-gradient of f at x_0. □

References

[1] E. Asplund, *Fréchet differentiability of convex functions*, Acta Math. **121** (1968), 31–47.

[2] J. W. Ellis, *A general set-separation theorem*, Duke Math. J. **19** (1952), 417–422.

[3] K. Fan, *On the Krein–Milman theorem*, Convexity, Proc. Sympos. Pure Math., vol. 7, Amer. Math. Soc., Providence, 1963, pp. 211–220.

[4] S. Mazur, *Über konvexe Menge in lineare normierte Räumen*, Studia Math. **4** (1933), 70–84.

[5] S. Mazur et S. Ulam, *Sur les transformations isométriques d'espaces vectoriels normés*, C. R. Acad. Sci. Paris **194** (1932), 946–948.

[6] K. Menger, *Untersuchungen über allgemeine Metrik I, II, III*, Math. Ann. **100** (1928), 75–165.

[7] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Math., vol. 1364, Springer-Verlag, Berlin, 1989.

[8] D. Preiss and L. Zajíček, *Stronger estimates of smallness of sets of Fréchet nondifferentiability of convex functions*, Proc. 11-th Winter School, Rend. Circ. Mat. Palermo (2) Suppl. **3** (1984), 219–223.

[9] S. Rolewicz, *On extension of Mazur theorem on Lipschitz functions*, Arch. Math. (Basel) **63** (1994), 535–540.

[10] ———, *On a globalization property*, Appl. Math. (Warsaw) **22** (1993), 69–73.

[11] ———, *On subdifferentials on nonconvex sets*, Dissertationes Math. **340** (1995), 301–308.

[12] V. P. Soltan, *Introduction to an Axiomatic Theory of Convexity*, Kishinev, 1984. (Russian)

Manuscript received February 20, 1995

S. Rolewicz
Institute of Mathematics
Polish Academy of Sciences
Śniadeckich 8
P.O. Box 137
00-950 Warszawa, POLAND

E-mail address: rolewicz@impan.gov.pl