Control and Cybernetics

vol. **23** (1994) No. 1/2

Convex analysis without linearity

by

Stefan Rolewicz*)

Institute of Mathematics
Polish Academy of Sciences
Warsaw
Poland

In the papaer a possibility of constructing convex analysis without linearity is presented.

1. Φ–convexity, Φ–subgradients and general duality

In 1943 F.Werfel, Werfel (1959), formulated the conjecture that there is a possibility of existence of anti–Semitism without Jews. Further developments showed that Werfel was right.

In the present paper we shall consider a similar problem, namely the possibility of existence of convex analysis without linearity.

It is obvious that in this case it is necessary to extend notions considered in convex analysis, such as convexity, subgradient, subdifferential in order for them to apply to this more general case.

The first step in this direction was done in 1963 by Ky Fan (Ky Fan, 1963), who introduced the notion of Φ–convexity. Let $(X, \|.\|)$ be a Banach space and let Φ be a family of continuous real valued functions defined on X. A set $A \subset X$ will be called Φ*–convex* set if for each $p \notin A$ there are $\phi_p \in \Phi$ and $c \in R$ such that

$$\phi_p(p) \geq c \tag{1.1}$$

and

$$\phi_p(x) < c \tag{1.2}$$

for all $x \in A$.

*) The paper was partially supported by the Polish Committe for Scientific Research under grant no. 2 2009 91 02

It is easy to see that a set A is Φ–convex if and only if there are a subfamily $\Phi_A \subset \Phi$ and constants $c_\phi, \phi \in \Phi_A$, such that the set A can be represented in the following way

$$A = \bigcap_{\phi \in \Phi_A} \{x \in X : \phi(x) \le c_\phi\}. \tag{1.3}$$

For a given set A by its *Φ–convexification* we shall call the smallest Φ–convex set B containing A. The Φ–convexification shall be denoted by $conv_\Phi A$. It is not difficult to check that

$$conv_\Phi A = \bigcap_{\phi \in \Phi} \{x \in X : \phi(x) \le sup_{y \in A} \phi(y)\}. \tag{1.4}$$

Ky Fan introduced a notion of Φ–extremal points (we shall not give this definition here since it will be not essential in further considerations) and he proved the following extension of the Krein–Milman theorem

THEOREM 1.1 (Ky Fan, 1963) *Let A be a compact Φ–convex set in a Banach space. Then A is Φ–convexification of its extremal points.*

The next essential push came from the optimization theory. Kurcyusz (1975), Dolecki–Kurcyusz (1978) and Balder (1977) extended the notion of duality to the non convex case.

Let an arbitrary set X, called later the space, be given. Let Φ be a family (a class) of functions defined on X and admitting values in $\overline{R} = R \cup \{-\infty\} \cup$ $\cup \{+\infty\}$.

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

$$f(x) - f(x_0) \ge \phi(x) - \phi(x_0). \tag{1.5}$$

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

Observe that the order in real number induces the order on real valued functions. We shall write $g \le f$ without writing the argument if $g(x) \le f(x)$ for all $x \in X$. For a given function f we shall take

$$f^\Phi(x) = sup\{\phi(x) + c : \phi \in \Phi, c \in R, \phi + c \le f\}. \tag{1.6}$$

The function $f^\Phi(x)$ is called *Φ–convexification* of the function f. If $f^\Phi(x) =$ $= f(x)$ we say that the function f is *Φ–convex*. The set of all Φ–convex functions will be denoted by Φ_{conv}.

Let

$$f^*(\phi) = -c_f(\phi) = sup_{x \in X}[\phi(x) - f(x)]$$

The function $f^*(\phi)$ will be called *Fenchel dual function* (or *Fenchel conjugate function*), since in the case when $(X, \|.\|)$ is a Banach space and $(X^*, \|.\|^*)$ is the conjugate space (i.e. the class of all continuous linear functionals defined on X) this notion was introduced by Fenchel (see Fenchel, 1949; and Fenchel, 1951). For non–linear ϕ it was investigated by Moreau, (Moreau, 1963; and Moreau, 1966) under the name of inf–convolutions.

PROPOSITION 1.2 *A function ϕ_0 is a Φ–subgradient of a function f at the point x_0 if and only if ϕ_0 and x_0 give equality in the Fenchel–Moreau inequality, i.e.*

$$f(x_0) + f^*(\phi_0) = \phi_0(x_0). \tag{1.7}$$

Observe that the space X induces on the family Φ also family of functions by the formula $x(\phi) = \phi(x)$. This family will also be denoted as X. Thus for functions defined on Φ we can speak of X–convexity.

PROPOSITION 1.3 *The Fenchel dual function $f^*(\phi)$ is X–convex.*

The following natural question arises. Can we determine the function $f(x)$ having its dual $f^*(\phi)$? The answer is in general negative, it is positive, though, for Φ–convex functions.

By the *second Fenchel dual* we shall call the Fenchel dual function to a dual to a Fenchel dual function and we shall denote it by $f^{**}(x)$.

THEOREM 1.4 (Balder,1977; Dolecki,Kurcyusz,1978; Elster,Neshe,1974; Kurcyusz;1975) *For arbitrary real valued function f the second Fenchel dual is equal to Φ–convexification of the function f,*

$$f^{**}(x) = f^{\Phi}(x). \tag{1.8}$$

2. Generalization of Mazur Theorem for Φ–convex functions

In 1933 Mazur proved the following

THEOREM 2.1 (Mazur,1933) *Let $(X, \|.\|)$ be a separable real Banach space. Let $f(x)$ be a real valued convex continuous function defined on an open convex subset $\Omega \subset X$. Then there is a subset A of the first category such that on the set $\Omega \setminus A$ the function f is Gateaux differentiable.*

The result of Mazur was a starting point for the theory of differentiability of convex functions (see for example the book Phelps,1989).

By simple observation of the convex functions of one variable we obtain that they are differentiable if and only if the subgradient at x_0 is unique. As a consequence we obtain that continuous convex function is Gateaux differentiable at x_0 if and only if the subgradient at x_0 is uniquely determined.

Having this in mind we can reformulate the Mazur theorem in the following way

THEOREM 2.1' *Let $(X, \|.\|)$ be a separable real Banach space. Let $f(x)$ be a real valued convex continuous function defined on an open convex subset $\Omega \subset X$. Then there is a subset A of the first category such that on the set $\Omega \setminus A$ the the subgradient of the function f is uniquely determined.*

In order to extend the Theorem 2.1' for the non–linear case we need a notion of monotone multifunction. A multifunction Γ mapping X into 2^{Φ} will be called *monotone* multifunction, if for $\phi_x \in \Gamma(x)$, $\phi_y \in \Gamma(y)$ we have

$$\phi_x(x) + \phi_y(y) - \phi_x(y) - \phi_y(x) \geq 0. \tag{2.1}$$

In a particular case, when X is a linear space, and Φ is a linear space consisting of linear functional, denoting $\phi(x)$ by $\phi(x) =< \phi, x >$ we can rewrite (2.1) in the classical form

$$< \phi_x - \phi_y, x - y > \geq 0. \tag{2.2}$$

As a trivial consequence of the definition we obtain that for a given function f the subdifferential $\partial^{\Phi} f|_x$ as a multifunction of x is a monotone multifunction.

It is interesting to know which conditions on the metric space (X, d) and on the class of real valued functions Φ warrant that for any monotone multifunction $\Gamma : X \to 2^{\Phi}$, there is a set A_{Γ} of the first category such that outside the set A_{Γ} the multifunction Γ is single valued. For this purpose we shall introduce some new notions.

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

Let

$$d_L(\phi_1, \phi_2) = sup_{x_1, x_2 \in X,\ x_1 \neq x_2} \frac{|[\phi_1(x_1) - \phi_2(x_1)] - [\phi_1(x_2) - \phi_2(x_2)]|}{d_X(x_1, x_2)} \tag{2.3}$$

It is easy to see that d_L is a quasimetric, i.e. it is symmetric and satisfies the triangle inequality. Observe that in $d_L(\phi_1, \phi_2) = 0$, then 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 Φ/R.

Let Φ be a family of Lipschitz functions. We assume that the family Φ is linear. 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

$$\phi(y) - \phi(x) \geq k d_L(\phi, 0)\, d_X(y, x) \tag{2.4}$$

we say that the family Φ has *monotonicity property* with the constant k. It is obvious that the linear continuous functionals over Banach space have monotonicity property with any constant k, $0 < k < 1$.

Having this notion and using the method of Preiss and Zajicek, (Preiss, Zajicek, 1984), we can obtain

THEOREM 2.2 (Rolewicz, 1994B) *Let (X, d_X) be a complete metric space. Let Φ be a linear family of Lipschitz functions having monotonicity property with a constant k. Assume that Φ/R is separable in the metric d_L. Let Γ be a monotone multifunction mapping X into Φ such that $\Gamma(x) \neq \emptyset$ for all $x \in X$. Then there exist a set A of the first category such that Γ is single valued and continuous on the set $X \setminus A$.*

Since the subdifferential $\partial f|_x$ is a monotone multifunction of x we trivially obtain

COROLLARY 2.3 *Let X be a complete metric space. Let Φ be a linear class of Lipschitz functions having monotonicity property with a certain constant k. Suppose that Φ is separable in the metric d_L. Let $f(x)$ be a Φ–convex function having at each point a Φ–subgradient. Then there is a set A of the first category such that outside the set A the subdifferential $\partial f|_x$ is single valued and it is 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 mononicity property with any constant smaller than 1. Thus we can formulate

COROLLARY 2.4 *Let X be a Banach space having separable dual X^*. Let $f(x)$ be a convex continuous function. Then there is a set A of the first category such that on the set dom $f \setminus A$ the subdifferential $\partial f|_x$ is single valued and it is continuous in the norm topology.*

Corollary 2.4 is a weak version of the Mazur theorem, Mazur (1933), (since in Mazur theorem the separability of the space X is requested only).

In the Corollary 2.3 we can weaken the assumption of monotonicity property with constant k by its local version. We say that a class Φ has *local monotonicity property* if for each $x \in X$ there is a neighbourhood U of x such that the family $\Phi|_U$ of the restriction of the family Φ to the set U has monotonicity property with a certain constant k_U. A family of locally Lipschitz functions Φ is called *locally separable*, if for each $x \in X$ there is a neighbourhood U of x such that the family $\Phi|_U$ of the restriction of the family Φ to the set U is separable in the Lipschitz metric.

Having these notions we can show

THEOREM 2.5 *Let (X, d_X) be a complete separable connected metric space. Let Φ be a linear family of locally Lipschitz functions having 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 exist a set A of the first category such that Γ is single valued and continuous on the set $X \setminus A$.*

The following interesting question arises : which classes Φ have monotonicity property with constant k? It is easy to see that if X is a convex subset of a normed space and Φ consist of linear functionals restricted to X, then Φ has monotonicity property with an arbitrary constant $0 < k < 1$. Similarly, if X is an open subset of a normed space and Φ consist of linear functionals restricted to X, then Φ has monotonicity property with an arbitrary constant $0 < k < 1$. If X is neither convex nor open the situation can be different. For example if X is a circle in R^2, then no one linear functional $\phi \in \Phi$, $\phi \neq 0$ has monotonicity property with any constant k. Thus we have a following problem: which sets in R^n (or more general in a Banach space) have this property that any family of linear functionals restricted to X has monotonicity property with a constant k?

3. Relation between uniform convexity of a given function and uniform smoothness of the conjugate

The next important step in convex analysis was done by Asplund, (Asplund, 1968), who proved a quantitative version of duality theory in the case of Banach spaces.

In his fundamental paper he proved the following

THEOREM 3.1 (Asplund,1968, see also Bronstedt,1964) *Let $f(x)$ be a lower-semicontinuous convex function defined on X. Let γ be a convex functions mapping the interval $[0, +\infty)$ into $[0, +\infty]$ such that $\gamma(0) = 0$. For a fixed $x_0 \in X$ and $x_0^* \in X^*$ the following inequalities are equivalent*

$$f(x) - f(x_0) \geq x_0^*(x - x_0) + \gamma(\|x - x_0\|) \text{ for all } x \in X, \tag{3.1}$$

$$f^*(x^*) - f(x_0^*) \leq (x^* - x_0^*)(x_0) + \gamma^*(\|x^* - x_0^*\|^*) \text{ for all } x^* \in X^*. \tag{3.2}$$

If (3.1) holds with reversed inequality sign, then (3.2) also reverses.

In Theorem 3.1 γ^* denotes the function conjugate to γ,

$$\gamma^*(t) = sup_{u>0}[ut - \gamma(u)]. \tag{3.3}$$

The Asplund theorem can be extended to the case of general duality for the case when Φ is a class of Lipschitz functions defined on a metric space X.

THEOREM 3.2 (Rolewicz, 1993A) *Let $f(x)$ be a Φ-convex function. Suppose that*

$$f^*(\phi) \geq f^*(\phi_0) + \phi(x_0) - \phi_0(x_0) + \gamma(d_L(\phi, \phi_0)) \tag{3.4}$$

holds. Then

$$f(x) \leq f(x_0) + \phi_0(x) - \phi_0(x_0) + \gamma^*(d(x, x_0)). \tag{3.5}$$

Let (X, d) be a metric space. Let Φ consist of Lipschitzian functions. As it was shown before the metric d induces on the space Φ/R a metric d_L. Observe that X can be interpreted as a set of Lipschitzian functions of $(\Phi/R, d_L)$. Then we can consider on the space X a corresponding Lipschitzian metric which we denote as $d_L(d_L(x, y))$.

Using this metric we can obtain the following proposition

THEOREM 3.3 (Rolewicz, 1993A) *Suppose that*

$$f(x) \geq f(x_0) + \phi_0(x) - \phi_0(x_0) + \gamma(d_L(d_L(x, x_0))). \tag{3.6}$$

Then

$$f^*(\phi) \leq f^*(\phi_0) + \phi(x_0) - \phi_0(x_0) + \gamma^*(d_L(\phi, \phi_0)), \tag{3.7}$$

In the case when the metric $d_L(d_L(x, y))$ coincides with the initial metric $d(x, y)$, $d_L(d_L(x, y)) = d(x, y)$, the formula (3.7) obtains a simpler form

$$f(x) - \phi_0(x) \geq f(x_0) - \phi_0(x_0) + \gamma(d(x, x_0)). \tag{3.7'}$$

THEOREM 3.4 (Rolewicz, 1993A) *Let (X, d) be a metric space. Let Φ denote a class of Lipschitzian functions defined on X, such that for each $x_0, \phi_0, \phi, t, \delta, \varepsilon > 0$ there is x such that*

$$|d(x, x_0) - t| < \delta t \tag{3.8}$$

and

$$|[\phi(x) - \phi(x_0)] - [\phi_0(x) - \phi_0(x_0)] - d_L(\phi, \phi_0) d(x, x_0)| < \varepsilon. \tag{3.9}$$

Let $f(x)$ be a Φ–convex function. If ϕ_0 is a Φ–subgradient of the function $f(x)$ at a point x_0 and

$$f(x) \leq f(x_0) + \phi_0(x) - \phi_0(x_0) + \gamma^*(d(x, x_0)), \tag{3.10}$$

then

$$f^*(\phi) - \phi(x_0) \geq f^*(\phi_0) - \phi_0(x_0) + \gamma(d_L(\phi, \phi_0)). \tag{3.11}$$

4. Globalization property

Now we shall consider a localizations of notions of Φ–convexity and Φ–subgradients.

We say that the function $f(x)$ is *locally Φ–convex* if for each $x_0 \in X$ there is a neighbourhood U of x_0 such that the function $f|_U(x)$ is $\Phi|_U$-convex, where $f|_U(x)$ and $\Phi|_U$ denote the restriction of the function $f(x)$ and the class Φ to the set U. Of course each Φ–convex function is locally Φ-convex, too. The converse is not true, Rolewicz (1993B).

A function $\phi \in \Phi$ is called a *local Φ–subgradient* of the function f at a point x_0 if there is a neighbourhood U of the point x_0 such that for all $x \in U$

$$f(x) - f(x_0) \geq \phi(x) - \phi(x_0) \tag{4.1}$$

holds.

It is easy to show that the fact of possessing of a local Φ–subgradient at each point does not imply that a function f has a Φ–subgradient at each point. Even more, the function f need not to be Φ–convex as follows from Rolewicz (1993B).

It is interesting, however, that there are classes Φ such that the existence of a local Φ–subgradient of a locally Φ–convex function $f(x)$ at each point $x_0 \in X$ implies the existence of a global Φ–subgradients of the function $f(x)$ at each point. If such a situation occurs we say that the family Φ has the *globalization property*. If each local Φ–subgradient can be extended to the global one we say that the family Φ has the *strong globalization property*.

If it holds for functions $f(x)$ satisfying the additional condition that there is $\phi \in \Phi$ such that

$$inf\ [f(x) - \phi(x)] > -\infty, \tag{4.2}$$

then we say that the family Φ has the *bounded globalization property* (resp. *bounded strong globalization property*).

We say that the set A has the *linear globalization property* if the family X^* restricted to A has the globalization property. We say that the set A has the *linear bounded globalization property* if the family X^* restricted to A has the bounded globalization property.

PROPOSITION 4.1 (Rolewicz, 1994A) *A closed set A has the strong linear globalization property if and only if it is convex.*

PROPOSITION 4.2 (Rolewicz, 1993B; Rolewicz, 1994A) *Let A be a boundary of a convex bounded open set B in a Banach space $(X, \|.\|)$, $A = Fr\ B$. Then the set A has the bounded linear globalization property.*

COROLLARY 4.3 (Rolewicz, 1993B; Rolewicz, 1994A) *Let A be a boundary of a convex bounded open set B in a finite dimensional Banach space $(X, \|.\|)$, $A = Fr\ B$. Then the set A has the linear globalization property.*

Without boundness of the set B Proposition 2 does not hold as follows from

EXAMPLE 4.4 *Let $X = R^2$ and let $A = \{(x, y) : |y| = 1\}$. It is easy to see that A is a boundary of an open convex set $B = \{(x, y) : |y| < 1\}$ and that the set A does not have the linear globalization property.*

The set A is not connected. As an example of connected set we can take a set $A_0 = Fr\ B_0$, where $B_0 = \{(x, y) : |y| < 1,\ x > 0\}$

PROPOSITION 4.5 *Let A be a closed set in a Banach space $(X, \|.\|)$. Let Φ be a restriction of linear functionals to A. If there are a point $p_0 \notin A$ and a vector v such that there are $t_1 < 0 < t_2 < t_3$ such that $p_0 + t_i v \in A$, $i = 1, 2, 3$, then the set A does not have bounded linear globalization property.*

COROLLARY 4.6 *Let A be a closed set in a Banach space $(X, \|.\|)$. If the set A has non-empty interior, Int $A \neq \emptyset$, then the set A has linear globalization property if and only if it is convex.*

COROLLARY 4.7 *Let A be a closed set in a Banach space $(X, \|.\|)$. Suppose that Int $A = \emptyset$ and that there is a closed set $B \subset A$, such that the set B is a boundary of an open set C, $B = Fr\ C$. Then the set A has bounded linear globalization property if and only if $A = B$ and the set C is convex.*

References

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

BALDER E.J. (1977) An extension of duality–stability relations to non–convex optimization problems, *SIAM Jour, Contr. Optim.* **15**, pp 329–343.

BRONSTEDT A. (1964) Conjugate convex functions in topological vector spaces, *Mat.–Fys. Medel Danska Vod Selsk* No **2**.

DOLECKI S. AND KURCYUSZ S. (1978) On Φ–convexity in extremal problems. *SIAM Jour. Control and Optim.* **16**, pp. 277–300, **224**, pp. 193–216.

ELSTER K.H. AND NEHSE R. (1974) Zur Theorie der Polarfunktionale, *Math. Operations forsch. und Stat. ser. Optimization* **5**, pp. 3–21.

KY FAN (1963) On the Krein–Milman theorem, *in Convexity, Proc. of Symp. in Pure Math.* **7** *Amer. Math. Soc. Providence*, pp. 211–220.

FENCHEL W. (1949) On conjugate convex functions, *Cand. Jour. Math.* **1**, pp.73–77.

FENCHEL W. (1951) *Convex cones, sets and functions* , Princeton Univ.

KURCYUSZ S. (1975) Some remarks on generalized Lagrangians, *Proc. 7–th IFIP Conference on Optimization Technique*, Nice, September 1975, Springer–Verlag.

MAZUR S. (1933) Über konvexe Menge in lineare normierte Raümen, *Stud. Math.* **4**, pp 70–84.

MOREAU J.J. (1963) Inf–convolutions des fonctions numérique sur un espace vectoriel, *C.R. Acad.Sc. Paris.* **256**, pp 5047–5049.

MOREAU J.J. (1966) Fonctionelles convexes, Seminaire sur les équations aux dérivees partielles, College de France, Paris.

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

PREISS D. AND ZAJICEK L. (1984) Stronger estimates of smallness of sets of Fréchet nondifferentiabilty of convex functions, *Proc. 11–th Winter School, Suppl. Rend. Circ. Mat di Palermo*, ser II, **3**, pp 219–223.

ROLEWICZ S. (1993A) Generalization of Asplund inequalities on Lipschitz functions, *Archiv der Math.* **61**, pp. 484–488.

ROLEWICZ S. (1993B) On a globalization property, *Appl. Math.* **22**.

ROLEWICZ S. (1994A) On subdifferentials on non-convex sets. *Ann. Pol. Math.* (in print).

ROLEWICZ S. (1994B) On extension of Mazur theorem on Lipschitz functions, *Arch. der Math.* (submitted).

WERFEL F. (1959) *Jacobowsky und der Oberst*, Drama II, pp. 241–340, S.Fisher Verlag, Frankfurt am Main.