HOW TO DEFINE ”CONVEX FUNCTIONS” ON DIFFERENTIABLE MANIFOLDS

Stefan Rolewicz

Institute of Mathematics of the Polish Academy of Sciences
Śniadeckich 8, 00–956 Warszawa, P.O. Box 21, Poland

e-mail: rolewicz@impan.gov.pl

Abstract

In the paper a class of families $\mathcal{F}(M)$ of functions defined on differentiable manifolds $M$ with the following properties:

1. if $M$ is a linear manifold, then $\mathcal{F}(M)$ contains convex functions,
2. $\mathcal{F}(\cdot)$ is invariant under diffeomorphisms,
3. each $f \in \mathcal{F}(M)$ is differentiable on a dense $G_\delta$-set,

is investigated.

Keywords: Fréchet differentiability, Gateaux differentiability, locally strongly paraconvex functions, $C^{1,\alpha}$-manifolds.

2000 Mathematics Subject Classification: 58C20, 46G05, 26E15.

Let $(X,\|\cdot\|)$ 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.,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

for all $x, y \in \Omega$ and $t$, $0 \leq t \leq 1$.

Mazur (1933) proved that in the case of a separable Banach space $X$ there is a dense $G_\delta$-subset $A_G$ such that on the set $A_G$ the function $f$ is Gateaux differentiable. Asplund (1968) showed that if in the dual space $X^*$ there exists an equivalent locally uniformly rotund norm, then there is a dense $G_\delta$-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. What is more, a space $X$ is an Asplund space if and only if each of its separable subspace $X_0 \subset X$ has a separable dual (Phelps (1989)).

The aim of this note is to obtain similar results for functions defined on differentiable manifolds. The first problem is how to define ”convex function” in this case. For this purpose we shall introduce a class of families $\mathcal{F}(M)$ of functions defined on differentiable manifold $M$ over a Banach space $E$ with the following properties:

1. If $M$ is a linear manifold, then $\mathcal{F}(M)$ contains convex functions,
2. $\mathcal{F}(-)$ is invariant under diffeomorphisms,
3. each $f \in \mathcal{F}(M)$ is
   a. Fréchet differentiable on a dense $G_\delta$-set provided $E$ is an Asplund space,
   b. Gateaux differentiable on dense $G_\delta$-set provided $E$ is separable.

At the beginning we recall the notion of differentiable manifolds.

Let $E, F$, be real Banach spaces. We say that a function $\psi : E \to F$ is of the class $C^{1,u}_{E,F}$ if it is continuously differentiable and, moreover, that differential $\partial \psi|_x$ is locally uniformly continuous as a function of $x$ in the norm topology. Of course, if $\psi \in C^{1,u}_{E,F}$, then $\psi$ belongs to the class of continuously differentiable functions, $\psi \in C^1_{E,F}$. The converse is true if $E$ is finite dimensional.

If $E = F$ we denote briefly $C^{1,u}_{E,E} = C^1_{E}$.

Now we shall determine $C^{1,u}_{E,E}$-manifold in the classical way (compare Lang (1962)).

Let $M$ be a set. A $C^{1,u}_E$-atlas is a collection of pairs $(U_i, \phi_i)$ ($i$ ranging in some indexing set) satisfying the following conditions:

AT 1. Each $U_i$ is a subset of $M$ and $\{U_i\}$ covers $M$,

AT 2. Each $\phi_i$ is a bijection of $U_i$ onto an open subset $\phi_i(U_i)$ of the space $E$, and for all $i, j$, $\phi_i(U_i \cap U_j)$ is an open subset of the space $E$,

AT 3. The map $\phi_j \phi_i^{-1}$ mapping $\phi_i(U_i \cap U_j)$ onto $\phi_j(U_i \cap U_j)$ is of the class $C^{1,u}_E$ for all $i, j$.

Each pair $(U_i, \phi_i)$ is called a chart. If $x \in U_i$, then the pair $(U_i, \phi_i)$ is called a chart at $x$.

Observe that AT 3 implies that $(\phi_j \phi_i^{-1})^{-1} = \phi_i \phi_j^{-1} \in C^{1,u}_E$. 
Suppose now that $M$ is a topological space and let $U$ be an open set in $M$. Suppose that there is a topological isomorphism $\phi$ mapping $U$ onto an open set $U' \in E$. We say that $(U, \phi)$ is compatible with the $C^{1,u}_E$-atlas $(U_i, \phi_i)$ if for all $i$ the maps $\phi_i \phi^{-1}$ and $\phi_i \phi_i^{-1}$ belong to $C^{1,u}_E$. We say that two $C^{1,u}_E$-atlases are compatible if each chart of the first is compatible with the other $C^{1,u}_E$-atlas.

A topological space $M$ equipped with $C^{1,u}_E$-atlas $(U_i, \phi_i)$ we shall call $C^{1,u}_E$-manifold.

Let $M$ be a $C^{1,u}_E$-manifold. Let $(U_i, \phi_i)$ be a $C^{1,u}_E$-atlas on $X$. Let $f(\cdot)$ be a real-valued function defined on $X$. We say that the function $f(\cdot)$ is Fréchet (Gateaux) differentiable at $x_0 \in U_i$ if the function $f(\phi_i^{-1}(\cdot))$ is Fréchet (resp. Gateaux) differentiable at $\phi_i(x_0)$. Since for every Fréchet differentiable at $\phi_i(x_0)$ function $g(\cdot)$ and any $\sigma(\cdot) \in C^{1,u}_E$ the function $g(\sigma(\cdot))$ is Fréchet differentiable at $\sigma(\phi_i(x_0))$, the definition of Fréchet differentiability is the same for all compatible $C^{1,u}_E$-atlases. Situation with Gateaux differentiability is not so nice. However, if we restrict ourselves to locally Lipschitz functions, the situation is the same, since for every locally Lipschitz Gateaux differentiable at $\phi_i(x_0)$ function $g(\cdot)$ and any $\sigma(\cdot) \in C^{1,u}_E$ the function $g(\sigma(\cdot))$ is Gateaux differentiable at $\sigma(\phi_i(x_0))$.

The problem how to define a "convex" function is much more difficult. It seems that a natural definition is as follows: we say that a function $f(\cdot)$ defined on $M$ is "convex" if $f(\phi_i^{-1}(\cdot))$ defined on $E$ is locally convex. This definition has a serious disadvantage. Namely, it is obvious that the "convexity" of the "convex functions" in this case ought be independent of the chart. In other words we ought to define a class $C$ of real-valued functions $f(\cdot)$ such that the domains of $f(\cdot)$ are open subsets $\text{dom} f = \Omega_f \subset E$ and

1. every locally convex function belongs to $C$,
2. if $f \in C$ and $\sigma(\cdot)$ is a local diffeomorphism of $\Omega_f$ then for each $x \in \Omega_f$, there is an open set $U$, $x \in U \subset \Omega_f$, such that $f_U(\cdot)$ being the restriction of $f(\sigma(\cdot))$ to the set $U$ belongs to $C$,
3. for each $f \in C$, the function $f(\cdot)$ is
   (a). Fréchet differentiable on a dense $G_\delta$-set of its domain provided $E$ is an Asplund space,
   (b). Gateaux differentiable on dense $G_\delta$-set of its domain provided $E$ is separable.

Having the class $C$ satisfying 1$C$ and 2$C$ and 3$C$, we can easily define the class...
of functions $\mathcal{F}(M)$ defined on manifolds and satisfying $1_{\mathcal{F}}$ and $2_{\mathcal{F}}$ and $3_{\mathcal{F}}$. Namely, we say that a function $f(\cdot)$ defined on a manifold $M$ with an $C^{1,\nu}_E$-atlas $(U_i, \phi_i)$ ($i$ ranging in some indexing set) belongs to $\mathcal{F}(M)$ if for all $i$ $f(\phi_i^{-1}(\cdot)) \in \mathcal{C}$.

The simplest example of the class $\mathcal{C}$ having properties $1_{\mathcal{C}}$ and $2_{\mathcal{C}}$ and $3_{\mathcal{C}}$ is the following class $\mathcal{C}_0$. We say that a function $f \in \mathcal{C}_0$, if for all $x \in \text{dom} f$ there are an open set $U$, $x \in U \subset \Omega_f$, a diffeomorphism $\sigma$ of $U$ onto $\sigma(U)$ and a locally convex function $g(\cdot)$ defined on $\sigma(U)$ and such that $f(\cdot) = g(\sigma(\cdot))$. It is easy to see that the class $\mathcal{C}_0$ has the requested property. In the case (b) we use the fact that locally convex function is locally Lipschitzian.

However, the class $\mathcal{C}_0$ has serious disadvantages. The first one is that there is not a nice description of this class similar to local convexity, the second is that the sum of two functions $f, g$ belonging to the class $\mathcal{C}_0$ and having the same domain may not belong to the class $\mathcal{C}_0$.

**Example 1.** Let $E = \mathbb{R}$. Let

$$f(x) = [\arctan(x - a)]^2$$

and

$$g(x) = [\arctan(x + a)]^2.$$

Of course, both functions $f, g \in \mathcal{C}_0$ as a composition of quadratic function and diffeomorphisms. Observe that for each $a$

$$f(a) + g(a) = f(-a) + g(-a) = [\arctan(2a)]^2 < \left(\frac{\pi}{2}\right)^2.$$

Let $a$ be chosen in such a way that $\arctan(a) > \frac{1}{\sqrt{2}} \left(\frac{\pi}{2}\right)$. Thus

$$f(0) + g(0) = 2[\arctan(a)]^2 > \sqrt{2} \left(\frac{\pi}{2}\right)^2.$$

It implies that $f(x) + g(x)$ has local strict maximum. Thus $f(\cdot) + g(\cdot) \notin \mathcal{C}_0$, since functions belonging to $\mathcal{C}_0$ do not have a maximum.

Of course we can replace $\mathcal{C}_0$ by its cone

$$\mathcal{C}_\infty = \left\{ f \mid f = \sum_{i=1}^{n} f_i(\cdot), \ f_i \in \mathcal{C}_0 \right\}.$$

It is easy to check that $\mathcal{C}_\infty$ has the requested property, but still there is no natural description of $\mathcal{C}_\infty$. 


In the paper we propose another class of functions, which seems more proper. It will be locally strongly paraconvex functions.

Now we recall the notion of strongly $\alpha(\cdot)$-paraconvex functions ([9]). Let $\alpha(\cdot)$ be a nondecreasing function mapping the interval $[0, +\infty)$ into the interval $[0, +\infty]$ such that

\begin{equation}
\lim_{t \to 0} \frac{\alpha(t)}{t} = 0.
\end{equation}

Let a real-valued continuous function $f(\cdot)$ be defined on an open convex subset $\Omega \subset X$. We say that the function $f(\cdot)$ is strongly $\alpha(\cdot)$-paraconvex if for all $x, y \in \Omega$ and $0 \leq t \leq 1$ we have

\begin{equation}
f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) + \min \{t, (1 - t)\} \alpha(\|x - y\|).
\end{equation}

The set of all strongly $\alpha(\cdot)$-paraconvex functions defined on $\Omega$ shall be denoted by $\alpha PC(\Omega)$. If there is an $\alpha(\cdot)$ satisfying (1) such that a function is strongly $\alpha(\cdot)$-paraconvex we say that it is strongly paraconvex. The set of all strongly paraconvex functions defined on $\Omega$ shall be denoted by $PC(\Omega)$.

Let $X$ be a real Banach space. Let $f(\cdot)$ be a real-valued function defined on an open subset $\Omega \subset X$. We say that $f(\cdot)$ is locally strongly paraconvex if for each $x_0 \in \Omega$ there is a convex open neighbourhood $U_{x_0}$ of $x_0$ such that the function $f(\cdot)$ restricted to $U_{x_0}$, $f\big|_{U_{x_0}}(\cdot)$, is strongly paraconvex.

The set of all locally strongly paraconvex functions defined on $\Omega$ shall be denoted by $PC^{\text{Loc}}(\Omega)$.

It is easy to see that the class $PC^{\text{Loc}}(\Omega)$ satisfies condition 1$\text{C}$.

The following proposition plays the essential role in showing that it also satisfies condition 2$\text{C}$.

**Proposition 2.** Let $\Omega_X$ ($\Omega_Y$) be an open convex set in a real Banach space $X$ (resp. $Y$). Let $\sigma$ be a mapping of $\Omega_X$ into $\Omega_Y$ such that the differentials of $\partial \sigma|_x$ are uniformly continuous functions of $x$ in the norm topology. Then there is a function $\beta(\cdot)$ mapping the interval $[0, +\infty)$ into the interval $[0, +\infty]$ such that

\begin{equation}
\lim_{t \to 0} \frac{\beta(t)}{t} = 0
\end{equation}
and such that for all \(x, y \in \Omega_X\) and \(0 \leq t \leq 1\)
\[
\|\sigma(tx + (1-t)y) - [t\sigma(x) + (1-t)\sigma(y)]\| \leq \min[t, (1-t)]\beta(||x - y||).
\]

**Proof.** We shall start the proof of Proposition 2 with a special case, namely when \(Y = \mathbb{R}\) is one dimensional. In other words, we consider a real-valued function \(f(\cdot)\) defined on an open convex set \(\Omega \subset X\). By our assumptions \(f(\cdot)\) is differentiable on \(\Omega\) and the differentials of \(f|_x\) are uniformly continuous functions of \(x\) in the norm topology. In other words, there is a function \(\beta_0\) mapping the interval \([0, +\infty)\) into the interval \([0, +\infty]\) such that
\[
\lim_{t \to 0} \beta_0(t) = 0
\]
and
\[
\|\partial f|_x - \partial f|_y\| \leq \beta_0(||x - y||).
\]

We define
\[
F(t) = f(tx + (1-t)y) - [tf(x) + (1-t)f(y)].
\]
It is easy to observe that \(F(0) = F(1) = 0\). Now we shall calculate its derivative
\[
\left.\frac{dF}{dt}\right|_t = \partial f\left|_{(tx + (1-t)y)}\right. (x - y) - f(x) + f(y).
\]
Since \(F(0) = F(1) = 0\), by the Rolle theorem there is \(t_0, 0 \leq t_0 \leq 1\), such that \(\frac{dF}{dt}|_{t_0} = 0\). Thus for arbitrary \(t, 0 \leq t \leq 1\)
\[
\left.\frac{dF}{dt}\right|_t \leq \left.\frac{dF}{dt}\right|_{t_0} \leq \left|\partial f\left|_{(tx + (1-t)y)}\right. - \partial f\left|_{(t_0x + (1-t_0)y)}\right.\right|(x - y)
\]
\[
\leq \beta_0\left(||(tx + (1-t)y) - (t_0x + (1-t_0)y)||\right)||x - y||
\]
\[
\leq \beta_0\left(||x - y||\right)||x - y|| = \beta\left(||x - y||\right),
\]
where the function \(\beta(t) = t/\beta_0(t)\) satisfies (1).
Since $F(0) = F(1) = 0$, by (6) we have
\[ F(t) = \int_0^t \frac{dF}{ds} \bigg|_s ds \leq t \beta\left(\|x - y\|\right) \]
and
\[ F(t) = \int_t^1 \frac{dF}{ds} \bigg|_s ds \leq (1 - t) \beta\left(\|x - y\|\right). \]
Therefore,
\[ F(t) \leq \min[t, (1-t)] \beta(\|x - y\|). \]

Now we consider the general case.

Since the differentials of $\partial \sigma_x$ are uniformly continuous functions of $x$ in the norm topology, there is a function $\beta_0$ mapping the interval $[0, +\infty)$ into the interval $[0, +\infty]$ satisfying (3) and
\[ \|\partial \sigma_x - \partial \sigma_y\| \leq \beta_0(\|x - y\|). \]
Take any functional $\phi \in Y^*$ of norm one. We define
\[ f_{\phi}(t) =: \phi\left(\sigma(tx + (1-t)y) - (t\sigma(x) + (1-t)\sigma(y))\right). \]
Observe that the differentials of the real-valued $f_{\phi}$, $\partial f_{\phi}\bigg|\frac{\partial \sigma}{\partial x}$ are uniformly continuous functions of $x$ in the norm topology. Thus by (7)
\[ f_{\phi}(t) \leq \min[t, (1-t)] \beta(\|x - y\|). \]
Since $\phi$ was an arbitrary linear functional of norm one by (10) we get
\[ \|\sigma(tx + (1-t)y) - t\sigma(x) + (1-t)\sigma(y)\| \]
\[ = \sup_{\{\phi : \|\phi\| = 1\}} \phi(\sigma(tx + (1-t)y) - t\sigma(x) + (1-t)\sigma(y)) \]
\[ = \sup_{\{\phi : \|\phi\| = 1\}} f_{\phi}(t) \leq \min[t, (1-t)] \beta(\|x - y\|). \]
By Proposition 2 we get:

**Theorem 3** ([16]). Let $\Omega_X$ (resp. $\Omega_Y$) be an open set in a real Banach space $X$ (resp. $Y$). Let $f(\cdot)$ be a real-valued locally strongly paracconvex function defined on $\Omega_Y$. Let $\sigma$ be a mapping of a $\Omega_X$ into $\Omega_Y$ such that the differentials of $\sigma_x$ are locally uniformly continuous functions of $x$ in the norm topology. Then the composed function $f(\sigma(\cdot))$ is locally strongly paraconvex.

**Proof.** Let $x_0 \in \Omega_X$. Since $f(\cdot)$ is a real-valued locally strongly paraconvex function, there are an open convex neighborhood of $\sigma(x_0)$ $U_{\sigma(x_0)} \subset \Omega_Y$ and a nondecreasing function $\alpha_U(\cdot)$ satisfying (1) such that for all $x, y \in U_{\sigma(x_0)}$ and $0 \leq t \leq 1$

\[(12) \quad f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) + \min[t, (1 - t)]\alpha_U(\|x - y\|).\]

Recall that $f(\cdot)$ restricted to $U_{\sigma(x_0)}$ is a Lipschitz function. We shall denote the corresponding Lipschitz constant by $M$. Thus by Proposition 1

\[(13) \quad \|f(\sigma(tx + (1 - t)y)) - f(t\sigma(x) + (1 - t)\sigma(y))\| \leq M\|\sigma(tx + (1 - t)y) - t\sigma(x) + (1 - t)\sigma(y)\| \leq M \min[t, (1 - t)]\beta(\|x - y\|).\]

Therefore,

\[f(\sigma(tx + (1 - t)y)) \leq f(t\sigma(x) + (1 - t)\sigma(y)) + M \min[t, (1 - t)]\beta(\|x - y\|) \leq tf(\sigma(x)) + (1 - t)f(\sigma(y)) + \min[t, (1 - t)]\alpha(\|\sigma(x) - \sigma(y)\|) \]

\[+ M \min[t, (1 - t)]\beta(\|x - y\|) = tf(\sigma(x)) + (1 - t)f(\sigma(y)) + \min[t, (1 - t)]\left(\alpha(\|\sigma(x) - \sigma(y)\|) + \beta(\|x - y\|)\right).\]
Since $\sigma(\cdot)$ is locally uniformly differentiable, it is also locally Lipschitz, i.e., there are a neighbourhood $V_{x_0}$ of $x_0$ and a constant $N$ such that for $x, y \in V_{x_0}$

$$\|\sigma(x) - \sigma(y)\| \leq N\|x - y\|.$$  

Let

$$\gamma(t) = \alpha(Nt) + \beta(t).$$

It is easy to check that $\gamma(\cdot)$ satisfies (1). Moreover by (14) and (15) the function $f(\sigma(\cdot))$ is strongly $\gamma(\cdot)$-paraconvex on $V_{x_0}$. Therefore it is locally strongly paraconvex. 

Condition 3c is an immediate consequence of

**Theorem 4** ([9–15, 17]). Let $\Omega_X$ be an open set in a real Banach space $X$. Let $f(\cdot)$ be a real-valued locally strongly paraconvex function defined on $\Omega_X$. Then the function $f(\cdot)$ is:

(a). Fréchet differentiable on a dense $G_\delta$-set provided $X$ is an Asplund space,

(b). Gateaux differentiable on dense $G_\delta$-set provided $X$ is separable.

Combining Theorems 3 and 4 we trivially get

**Theorem 5**. (Rolewicz (2007)). Let $\Omega_X (\Omega_Y)$ be an open set in a real Banach space $X$ (resp. $Y$). Let $f(\cdot)$ be a real-valued locally uniformly approximately convex function defined on $\Omega_Y$. Let $\sigma$ be a mapping of a $\Omega_X$ into $\Omega_Y$ such that the differentials of $\sigma|_x$ are locally strongly paraconvex function of $x$. Then the composed function $f(\sigma(\cdot))$ is:

(a). Fréchet differentiable on a dense $G_\delta$-set provided $X$ is an Asplund space,

(b). Gateaux differentiable on dense $G_\delta$-set provided $X$ is separable.

We say that a real-valued function $f(\cdot)$ defined on a $C^1_{\mathcal{E}}$-manifold $M$ is locally strongly paraconvex on $M$ if there is a $C^1_{\mathcal{E}}$-atlas $(U_i, \phi_i)$ such that for all $i$ the function $f(\phi_i^{-1}(\cdot))$ is locally strongly paraconvex on the set $\phi_i(U_i) \subset \mathcal{E}$.

Basing on Theorem 5 and the definitions of differentiability of functions on manifold we immediately obtain

**Theorem 6** ([16]). Let $M$ be a $C^1_{\mathcal{E}}$-manifold. Let $f(\cdot)$ be a real-valued locally strongly paraconvex function defined on $M$. Then it is:
S. Rolewicz

(a) Fréchet differentiable on a dense $G_δ$-set provided $E$ is an Asplund space, (b) Gateaux differentiable on a dense $G_δ$-set provided $E$ is separable.

Now we shall determine $C^{1,u}_E$-submanifold in the classical way (compare Lang (1962)).

Let $M$ be a $C^{1,u}_E$-manifold. Let $N$ be a subset of $M$. We assume that for each point $y \in N$ there exists a chart $(V, \psi)$ in $M$ such that $V_1 = \psi(V \cap N)$ is an open set in some Banach subspace $E_1 \subset E$. The map $\psi$ induces a bijection

$$\psi_1 : Y \cap V \to V_1$$

and moreover $\psi_1 \in C^{1,u}_{E_1}$.

The collection of pairs $(N \cap V, \psi_1)$ obtained in the above manner constitute the atlas for $N$. We shall call $N C^{1,u}_{E_1}$-submanifold of $M$.

**Theorem 7** ([16]). Let $M$ be a $C^{1,u}_E$-manifold. Let $N$ be its $C^{1,u}_{E_1}$-submanifold. Let $f(\cdot)$ be a real-valued locally strongly paraconvex function defined on $M$. Then the restriction $f \big|_N$ is locally strongly paraconvex function defined on $N$.

**Corollary 8.** Let $f(\cdot)$ be a convex function defined on $\mathbb{R}^n$. Let $M$ be an $m$-dimensional manifold, $m < n$, imbedded in $\mathbb{R}^n$. Then the restriction of the function $f(\cdot)$ to $M$ is differentiable on a dense $G_δ$-set.

**References**


How to define "convex functions" on …


Received 11 February 2009