On the points of non-differentiability of convex functions

David Pavlica

Abstract. We characterize sets of non-differentiability points of convex functions on $\mathbb{R}^n$. This completes (in $\mathbb{R}^n$) the result by Zajíček [2] which gives a characterization of the magnitude of these sets.

Keywords: differentiability, convex functions

Classification: Primary 26B05; Secondary 26B25

In the present paper we give a complete characterization of sets of non-differentiability points of convex functions on $\mathbb{R}^n$. For a convex function f on $\mathbb{R}^n$, $0 \leq k \leq n$, $S_k(f)$ is the set of all $x \in \mathbb{R}^n$ for which $\dim \partial f(x) \geq n-k$ ($\partial f(x)$ denotes the subdifferential of f at the point x). In [2] the following characterization of the magnitude of $S_k(f)$ is given.

Definition 1. A set $S \subset \mathbb{R}^n$ is called a δ-convex surface of dimension k ($k = 1, \dots, n-1$) if there exists a permutation π of the numbers $1, 2, \dots, n$ and $2n-2k$ convex functions $f_{k+1}, g_{k+1}, \dots, f_n, g_n$ defined on the whole space $\mathbb{R}^k$ such that

$$S = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_{\pi(j)} = f_j(x_{\pi(1)}, \dots, x_{\pi(k)}) - g_j(x_{\pi(1)}, \dots, x_{\pi(k)}) \\ \text{for } j = k+1, \dots, n\}.$$

Theorem Z. *A set $M \subset \mathbb{R}^n$ is a subset of the set $S_k(f)$ $(1 \leq k \leq n-1)$ for some convex function f defined on $\mathbb{R}^n$ iff M can be covered by countably many δ-convex surfaces of dimension k.*

It is known that, for any convex function $f\colon \mathbb{R}^n \to \mathbb{R}$, $S_k(f)$ is a F_σ-set. We shall prove the following theorem.

Theorem. *Let $1 \leq k \leq n-1$, P be an F_σ-subset of a countable union of δ-convex surfaces of dimension k. Then there exists a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $S_k(f) = P$ and f is differentiable at all points of $\mathbb{R}^n \setminus P$.*

In the proof we shall use the notion of a dual convex function.

The author was supported by the grants GAČR 201/00/0767 and GAČR 201/03/0931.

Definition. Let $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a convex function. The dual function f^* of the function f is defined on $(\mathbb{R}^n)^*$ by

$$f^*(x^*) = \sup_{x\in\mathbb{R}^n} \left(\langle x, x^*\rangle - f(x)\right), \qquad x^* \in (\mathbb{R}^n)^*.$$

It follows immediately from the definition that if $f, g\colon \mathbb{R}^n \to \mathbb{R}$ are convex functions, $f \le g$ and f^* is finite everywhere then g^* is finite everywhere.

As usual, we identify the dual space $(\mathbb{R}^n)^*$ with $\mathbb{R}^n$ and $\langle\cdot,\cdot\rangle$ denotes both duality and scalar product.

Facts. *If $f\colon \mathbb{R}^n \to \mathbb{R}$ is a convex function then*

(1) $(f^*)^* = f$,
(2) $x^* \in \partial f(x) \Leftrightarrow x \in \partial f^*(x^*)$,
(3) *if f^* is finite on $\mathbb{R}^n$, then the epigraph of f contains no non-vertical halflines.*

The statement (1) can be found in [1, Theorem 12.2], (2) in [1, Theorem 23.5] and (3) in [1, Corollary 13.3.1].

Fact (4)**.** *A closed convex set in $\mathbb{R}^n$ containing no halflines is bounded.*

Fact (4) can be easily proved by a compactness argument.

Fact (5)**.** *If f^* is finite on $\mathbb{R}^n$, then for each affine functional π, the set $\{x \in \mathbb{R}^n : f(x) \le \pi(x)\}$ is bounded.*

Fact (5) is a consequence of Facts (3) and (4).

If a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ is not differentiable at some point x then there exist $x^* \ne y^*$, $x^*, y^* \in \partial f(x)$, and therefore, by the fact (2), $x \in \partial f^*(x^*) \cap \partial f^*(y^*)$. Consequently there is a line segment on the graph of f^* with endpoints $(x^*, f^*(x^*)), (y^*, f^*(y^*))$. Conversely, if there is a line segment on the graph of f^* with a supporting linear functional $\langle x, \cdot\rangle$ (it means that for some $\alpha \in \mathbb{R}$ the graph of $\langle x, \cdot\rangle + \alpha$ contains this line segment and $\langle x, \cdot\rangle + \alpha \le f^*$) then f is not differentiable at x.

In particular, the dual function of a strictly convex function is differentiable everywhere.

In the proof of our theorem we need the following simple lemma.

Lemma 1. *Let T be a compact convex set in $\mathbb{R}^n$ with a non-empty interior, $h\colon T \to \mathbb{R}$ a convex function, $h|_{\partial T} \equiv 0$ and $h(x) < 0$ for some $x \in T$. Then there exists a convex function $\bar{h}\colon T \to \mathbb{R}$ such that $\bar{h}|_{\partial T} \equiv 0$, $\bar{h} \ge h$ on T and $\bar{h}$ is affine on no line segment in $\operatorname{int} T$.*

Proof: For a compact convex set C in $\mathbb{R}^n$ such that $0 \in \operatorname{int} C$, denote

$$\gamma(y|C) := \inf\{\mu \ge 0 : y \in \mu C\}, \quad y \in \mathbb{R}^n.$$

By [1, §15] $\gamma(\cdot|C)$ is a convex function (therefore it is continuous), obviously it is positively homogenous and equal to 1 on ∂C.

Let us denote for $x \in \operatorname{int} T$

$$h_x(z) := -h(x)\left(\gamma(z-x|T-x)-1\right), \quad z \in \mathbb{R}^n.$$

For $x \neq z$ denote $r_x(z)$ the point of intersection of ∂T and the halfline starting at x and containing z. It is easy to check that

$$r_z(y) = z + \frac{y-z}{\gamma(y-z|T-z)}, \quad z \in \operatorname{int} T,\ y \in \mathbb{R}^n \setminus \{z\}.$$

For $y = 2z - x$ we get

$$r_x(z) = r_z(y) = z + \frac{z-x}{\gamma(z-x|T-z)}, \quad x, z \in \operatorname{int} T,\ x \neq z.$$

Hence, for $z \in \operatorname{int} T$, $g(x) = r_x(z)$ is a continuous mapping on $\operatorname{int} T \setminus \{z\}$.

Clearly h_x is convex, $h_x \equiv 0$ on ∂T, $h_x < 0$ on $\operatorname{int} T$, $h_x \geq h$ on T, and h_x is affine on every halfline starting at the point x.

If $y \neq x \neq z$ and h_x is affine on $\operatorname{conv}\{y, z\}$ then it is affine on $\operatorname{conv}\{x, r_x(y), r_x(z)\}$ and therefore $\operatorname{conv}\{r_x(y), r_x(z)\} \subset \partial T$.

We choose a countable dense set $x_1, x_2, \ldots \in \operatorname{int} T$ and set

$$\bar{h} := \sum_{i=1}^{\infty} \frac{h_{x_i}}{2^i}.$$

Then obviously $\bar{h} \geq h$ on T and $\bar{h}|_{\partial T} \equiv 0$.

For a contradiction let us suppose $\bar{h}$ is affine on some line segment $\operatorname{conv}\{y, z\}$, $y \neq z$, $y, z \in \operatorname{int} T$. Then, for each i, h_{x_i} is affine on $\operatorname{conv}\{y, z\}$. We choose a sequence $\{x_{k_i}\}$ such that $x_{k_i} \to \frac{y+z}{2}$ for $i \to \infty$. Then we have

$$\operatorname{conv}\left\{r_{x_{k_i}}(y), r_{x_{k_i}}(z)\right\} \subset \partial T.$$

Letting $i \to \infty$ we get (since $g(x) = r_x(z)$ is a continuous mapping)

$$\operatorname{conv}\left\{r_{\frac{y+z}{2}}(y), r_{\frac{y+z}{2}}(z)\right\} \subset \partial T,$$

a contradiction. $\square$

Lemma 2. *Assume $F \subset \mathbb{R}^n$ is a closed subset of a δ-convex surface S of dimension k, $0 < k < n$. Then there exists a convex function $H\colon \mathbb{R}^n \to \mathbb{R}$ such that H is differentiable at all points of $\mathbb{R}^n \setminus F$ and $S_k(H) = F$.*

Proof: By Theorem Z there is a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $S \subset S_k(f)$. We may assume that f is strictly convex and f^* is finite everywhere since otherwise we take $f(x) + \|x\|^2$ (there exists an affine functional p such that $p \le f$ and since $(p(x) + \|x\|^2)^*$ is finite everywhere we have that $(f(x) + \|x\|^2)^*$ is finite everywhere too).

Therefore f^* is differentiable everywhere. Let us denote

$$F^* := \{x \in \mathbb{R}^n : \nabla(f^*)(x) \in F\}.$$

Since the mapping $\nabla(f^*)$ is continuous, F^* is closed. For $x \in \mathbb{R}^n$ denote by

$$p_x(z) = \langle z, x\rangle + \alpha_x$$

the supporting affine functional to f^* (it exists for all x since $(f^*)^* = f$ is finite everywhere). For $\varepsilon > 0$ let us denote

$$U_{x,\varepsilon} := \{z \in \mathbb{R}^n : f^*(z) < p_x(z) + \varepsilon\},$$
$$T_{x,\varepsilon} := \{z \in \mathbb{R}^n : f^*(z) \le p_x(z) + \varepsilon\}.$$

By the fact (5) applied to f^*, the set $T_{x,\varepsilon}$ is compact and clearly it is convex. The set $U_{x,\varepsilon}$ is open.

Claim. *For each $x \in \mathbb{R}^n \setminus F$,*

$$\lim_{\varepsilon \to 0_+} \operatorname{dist}(T_{x,\varepsilon}, F^*) > 0$$

holds.

Proof of Claim: Let us denote

$$W_x := \{z \in \mathbb{R}^n : f^*(z) = p_x(z)\} = \bigcap_{\varepsilon > 0} T_{x,\varepsilon}.$$

Clearly $W_x \cap F^* = \bigcap_{\varepsilon>0}(T_{x,\varepsilon} \cap F^*) = \emptyset$. Since $T_{x,\varepsilon} \cap F^*$ are compact, for some $\varepsilon_0 > 0$ we have $T_{x,\varepsilon_0} \cap F^* = \emptyset$. Thus $\operatorname{dist}(T_{x,\varepsilon_0}, F^*) > 0$ and consequently, since $g(\varepsilon) = \operatorname{dist}(T_{x,\varepsilon}, F^*)$ is a non-increasing function, our Claim is proved. $\square$

By above Claim we can, for every $x \in \mathbb{R}^n \setminus F$, fix $0 < \varepsilon_x < 1$ such that

$$\left[\operatorname{dist}(T_{x,\varepsilon_x}, F^*)\right]^2 \ge \varepsilon_x.$$

We have

$$\mathbb{R}^n \setminus F^* = \bigcup_{x \in \mathbb{R}^n \setminus F} U_{x,\varepsilon_x},$$

since, for $x^* \in \mathbb{R}^n \setminus F^*$, we have $x^* \in W_x \subset U_x$ for $x = \nabla f^*(x^*) \notin F$. Therefore there exist points $x_1, x_2, \ldots \in \mathbb{R}^n \setminus F$ such that

$$\mathbb{R}^n \setminus F^* = \bigcup_{i=1}^{\infty} U_{x_i,\varepsilon_{x_i}}.$$

According to Lemma 1, choose for each $i \in \mathbb{N}$ a convex function $h_i \colon T_{x_i,\varepsilon_{x_i}} \to \mathbb{R}$ such that

$$h_i|_{\partial T_{x_i,\varepsilon_{x_i}}} \equiv 0,$$

h_i is affine on no line segment in $U_{x_i,\varepsilon_{x_i}}$ and $h_i \geq f^* - p_{x_i} - \varepsilon_{x_i}$. Let us define

$$\begin{aligned} &\tilde{h}_i \colon \mathbb{R}^n \to \mathbb{R}, \\ &\tilde{h}_i = h_i + p_{x_i} + \varepsilon_{x_i} \qquad \text{on } \ T_{x_i,\varepsilon_{x_i}}, \\ &\quad\; = f^* \qquad\qquad\qquad\;\; \text{on } \ \mathbb{R}^n \setminus T_{x_i,\varepsilon_{x_i}}. \end{aligned}$$

Then $f^* \leq \tilde{h}_i \leq f^* + \varepsilon_{x_i}$.

Observation. *If h is a convex function on $\mathbb{R}^n$, $\bar{h}$ is a convex function on a compact convex set $T \subset \mathbb{R}^n$ and $\bar{h}|_{\partial T} \equiv h|_{\partial T}$, $\bar{h} \geq h$ on T, then the function*

$$\begin{aligned} \tilde{h} &= h \qquad \textit{on } \ \mathbb{R}^n \setminus T; \\ \tilde{h} &= \bar{h} \qquad \textit{on } \ T \end{aligned}$$

is convex.

Proof of Observation: For $n = 1$ it is elementary and the higher dimensional case is an immediate consequence of the 1-dimensional one. □

By this Observation functions $\tilde{h}_i$ are convex. Set

$$\tilde{h} := \sum_{i=1}^{\infty} \frac{\tilde{h}_i}{2^i}\,.$$

Clearly $\tilde{h} = f^*$ on F^*, and $0 \leq \tilde{h} - f^* \leq 1$. Hence $\tilde{h} < +\infty$. Moreover $\tilde{h}$ is affine on no line segment in $\mathbb{R}^n \setminus F^*$. Now we shall prove that $H := (\tilde{h})^*$ fulfills the assertion of the lemma. The function H is finite everywhere since $\tilde{h} \geq f^*$ and $(f^*)^*$ is finite everywhere.

Let $x \in F$. There exist affine independent $y_i \in \partial f(x)$, $i = 1, \dots, n-k+1$. By Fact (2) we have $x \in \partial f^*(y_i)$ and so $y_i \in F^*$, $i = 1, \dots, n-k+1$. Thus $\tilde{h}(y_i) = f^*(y_i)$ and consequently, since $\tilde{h} \geq f^*$, we have $x \in \partial\tilde{h}(y_i)$. Therefore $y_i \in \partial H(x)$, and so $x \in S_k(H)$.

Let us suppose for a contradiction that H is not differentiable at a point $x \notin F$. Then there exist $z_1 \neq z_2$, $z_1, z_2 \in \partial H(x)$. Thus $x \in \partial\tilde{h}(z_1) \cap \partial\tilde{h}(z_2)$. Further, $\tilde{h}$ is affine on no line segment in $\mathbb{R}^n \setminus F^*$, therefore $z_1, z_2 \in F^*$.

For each $i \in \mathbb{N}$ we have $f^* \leq \tilde{h}_i \leq f^* + \varepsilon_{x_i}$ and

$$\varepsilon_{x_i} \leq \Big[\operatorname{dist}(z_1, T_{x_i,\varepsilon_{x_i}})\Big]^2 .$$

Therefore

$$|f^*(z) - \tilde{h}_i(z)| \leq \varepsilon_{x_i} \leq \|z - z_1\|^2 \quad \text{for} \quad z \in T_{x_i,\varepsilon_{x_i}}.$$

Since also $f^*(z) = \tilde{h}_i(z)$ for $z \notin T_{x_i,\varepsilon_{x_i}}$, we have for all z

$$\begin{aligned} |f^*(z) - \tilde{h}(z)| &= \left|\sum_{i=1}^{\infty} \frac{1}{2^i}(f^*(z) - \tilde{h}_i(z))\right| \\ &\leq \sum_{i=1}^{\infty} \frac{1}{2^i}\|z - z_1\|^2 \leq \|z - z_1\|^2. \end{aligned}$$

This easily implies $\partial\tilde{h}(z_1) = \partial f^*(z_1)$, a contradiction with $x \in \partial\tilde{h}(z_1)$, $\partial f^*(z_1) \subset F$.

Lemma 3. *If $1 \leq k \leq n-1$ and $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots$, are convex functions, each differentiable at all points of $\mathbb{R}^n \setminus S_k(f_i)$, then there exists a convex function $f \colon \mathbb{R}^n \to \mathbb{R}$ such that*

$$S_k(f) = \bigcup_{i=1}^{\infty} S_k(f_i)$$

and f is differentiable at all points of $\mathbb{R}^n \setminus S_k(f)$.

Proof: Let us denote $B(0, r) := \{z : \|z\| \leq r\}$.

Choose $c_i > 0$, $i = 1, 2, \dots$, such that

$$|c_i f_i| \leq \frac{1}{2^i} \text{ on } B(0, i),$$

$$c_i f_i \text{ is Lipschitz with the constant } \frac{1}{2^i} \text{ on } B(0, i).$$

Set $f := \sum_{i=1}^{\infty} c_i f_i$. Clearly $S_k(f) \supseteq \bigcup_{i=1}^{\infty} S_k(f_i)$. Let us suppose for a contradiction f is not differentiable at some $x \in \mathbb{R}^n$ and all f_i are differentiable at x.

There exists $v \in \mathbb{R}^n$ such that $\|v\| = 1$ and

$$d := d^+ f(x)(v) + d^+ f(x)(-v) > 0,$$

where $d^+ f(x)(v) := \lim_{\lambda \to 0_+} \frac{f(x+\lambda v) - f(x)}{\lambda}$.

Find $j \in \mathbb{N}$ such that $2^{-j+1} < d$ and $x \in B(0,j)$. Since $\sum_{i=1}^{j} c_i f_i$ is differentiable at x,

$$d^+ \Big(\sum_{i=1}^{j} c_i f_i \Big)(x)(v) + d^+ \Big(\sum_{i=1}^{j} c_i f_i \Big)(x)(-v) = 0.$$

Further, $\sum_{i=j+1}^{\infty} c_i f_i$ is Lipschitz with the constant $\frac{1}{2^j}$ on $B(0,j+1)$, and therefore

$$d^+ \Big(\sum_{i=j+1}^{\infty} c_i f_i \Big)(x)(v) \leq \frac{1}{2^j}\,,$$
$$d^+ \Big(\sum_{i=j+1}^{\infty} c_i f_i \Big)(x)(-v) \leq \frac{1}{2^j}\,.$$

Thus we have $d^+ f(x)(v) + d^+ f(x)(-v) \leq \frac{1}{2^j} + \frac{1}{2^j} < d$, a contradiction. □

Proof of Theorem: Let $P = \bigcup_{i=1}^{\infty} F_i \subset \bigcup_{i=1}^{\infty} S_i$, where F_i is closed, S_i is a δ-convex surface of dimension k for all $i \in \mathbb{N}$. We have $P = \bigcup_{i,j=1}^{\infty} (F_i \cap S_j)$ and, since S_j are closed sets, we get by Lemma 2 functions $f_{i,j}$ differentiable at all points of $\mathbb{R}^n \setminus (F_i \cap S_j)$ such that $S_k(f_{i,j}) = F_i \cap S_j$. By Lemma 3 we then get a convex function f differentiable at all points of $\mathbb{R}^n \setminus P$ such that $S_k(f) = P$. □

Corollary. *Let $F \subset \mathbb{R}^n$, $1 \leq k \leq n-1$. Then $F = S_k(f)$ holds for some convex function f on $\mathbb{R}^n$ iff F is an F_σ-subset of a countable union of δ-convex surfaces of dimension k.*

Proof: By our Theorem, for every F_σ-subset P of a countable union of δ-convex surfaces of dimension k, there exists a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $S_k(f) = P$.

Conversely, for a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$, according to Theorem Z, $S_k(f)$ can be covered by countably many δ-convex surfaces of dimension k. And it is known that $S_k(f)$ is an F_σ-set. Since I do not know any reference to this simple result, I shall sketch the proof. Let $S_{k,j}(f)$ be the set of all points x such that there exist $u_0, \dots, u_k \in \partial f(x)$ such that $(u_i - u_0) \cdot (u_j - u_0) = 0$, $\|u_i - u_0\| = 1/j$ for all $i, j \in \{1, \dots, k\}$. Then we have $S_k(f) = \bigcup_{j=1}^{\infty} S_{k,j}(f)$ and $S_{k,j}(f)$ are closed sets. Thus we are done. □

References

[1] Rockafellar R.T., *Convex Analysis*, Princeton University Press, 1970.
[2] Zajíček L., *On the differentiability of convex functions in finite and infinite dimensional spaces*, Czechoslovak Math. J. **29** (104) (1979), 340–348.

Department of Mathematical Analysis, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague 8, Czech Republic

(*Received* October 2, 2003, *revised* January 12, 2004)