TUCS

Marko M. Mäkelä | Napsu Karmitsa | Ville-Pekka Eronen

On Generalized Pseudo- and Quasiconvexities for Nonsmooth Functions

Turku Centre for Computer Science

TUCS Technical Report
No 989, December 2010

On Generalized Pseudo- and Quasiconvexities for Nonsmooth Functions

Marko M. Mäkelä
University of Turku, Department of Mathematics
FI-20014 Turku, Finland
makela@utu.fi
Napsu Karmitsa
University of Turku, Department of Mathematics
FI-20014 Turku, Finland
napsu@karmitsa.fi
Ville-Pekka Eronen
University of Turku, Department of Mathematics
FI-20014 Turku, Finland
vpoero@utu.fi

TUCS Technical Report
No 989, December 2010

Abstract

Convexity plays a crucial role in mathematical optimization theory. In order to extend the existing results depending on convexity, numerous attempts of generalizing the concept have been published during years. Different types of generalized convexities has proved to be the main tool when constructing optimality conditions, particularly sufficient conditions for optimality.

The purpose of this paper is to analyze the properties of the generalized pseudo- and quasiconvexities for nondifferentiable locally Lipschitz continuous functions. The treatment is based on the Clarke subdifferentials and generalized directional derivatives.

Keywords: Generalized convexity; Clarke derivatives; Nonsmooth analysis

TUCS Laboratory
TUCS Laboratory

1 Introduction

Convexity plays a crucial role in mathematical optimization theory. Especially, in duality theory and in constructing optimality conditions convexity has been the most important concept since the basic reference by Rockafellar [15] was published. Recently there have been numerous attempts to generalize the concept of convexity in order to weaken the assumptions of the attained results (see e.g., [1, 2, 5, 9, 16, 18]). For an excellent survey of generalized convexities we refer to [14].

Generalized convexities have proved to be the main tool when constructing optimality conditions, particularly sufficient conditions. There exist a wide amount of papers published for smooth single-objective case (see [14] and references therein). For nonsmooth and multiobjective problems necessary conditions were derived for instance in [12, 13, 17].

In this paper, we analyze the properties of the generalized pseudo- and quasiconvexities for nondifferentiable locally Lipschitz continuous functions. The treatment is based on the Clarke subdifferentials and generalized directional derivatives [4]. The paper is organized as follows. In Section 2 we recall some basic tools from nonsmooth analysis. Sections 3 and 4 are devoted to generalized pseudo- and quasiconvexity, respectively. Also, some relations between generalized pseudo- and quasiconvexities are considered in Section 4. Finally, the derived results are summarized in Section 5.

2 Nonsmooth Analysis

In this section we collect some notions and results from nonsmooth analysis. Nevertheless, we start by recalling the notion of convexity and Lipschitz continuity. The function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$ we have

$$f\big(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}).$$

In what follows the considered functions are assumed to be locally Lipschitz continuous. A function is *locally Lipschitz continuous at a point* $\boldsymbol{x} \in \mathbb{R}^n$ if there exist scalars $K > 0$ and $\delta > 0$ such that

$$|f(\boldsymbol{y}) - f(\boldsymbol{z})| \leq K\|\boldsymbol{y} - \boldsymbol{z}\| \quad \text{for all } \boldsymbol{y}, \boldsymbol{z} \in B(\boldsymbol{x}; \delta),$$

where $B(\boldsymbol{x}; \delta) \subset \mathbb{R}^n$ is an open ball with center $\boldsymbol{x}$ and radius δ. Function is said to be *locally Lipschitz continuous on a set* $U \subseteq \mathbb{R}^n$ if it is locally Lipschitz continuous at every point belonging to the set U. Furthermore, if $U = \mathbb{R}^n$ the function is called *locally Lipschitz continuous*. Note that both convex and smooth (continuously differentiable) functions are always locally Lipschitz continuous (see, e.g., [4]).

Function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *Lipschitz continuous on a set* $U \subseteq \mathbb{R}^n$ if there exists a scalar K such that

$$|f(\boldsymbol{y}) - f(\boldsymbol{z})| \leq K\|\boldsymbol{y} - \boldsymbol{z}\| \quad \text{for all } \boldsymbol{y}, \boldsymbol{z} \in U.$$

If $U = \mathbb{R}^n$ then f is said to be *Lipschitz continuous*.

DEFINITION 2.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x} \in S \subseteq \mathbb{R}^n$. The *Clarke generalized directional derivative* of f at $\boldsymbol{x}$ in the direction of $\boldsymbol{d} \in \mathbb{R}^n$ is defined by

$$f^\circ(\boldsymbol{x}; \boldsymbol{d}) := \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \downarrow 0}} \frac{f(\boldsymbol{y} + t\boldsymbol{d}) - f(\boldsymbol{y})}{t}$$

and the *Clarke subdifferential* of f at $\boldsymbol{x}$ by

$$\partial f(\boldsymbol{x}) := \{\boldsymbol{\xi} \in \mathbb{R}^n \mid f^\circ(\boldsymbol{x}; \boldsymbol{d}) \geq \boldsymbol{\xi}^T \boldsymbol{d} \text{ for all } \boldsymbol{d} \in \mathbb{R}^n\}.$$

Each element $\boldsymbol{\xi} \in \partial f(\boldsymbol{x})$ is called a *subgradient* of f at $\boldsymbol{x}$.

Note that the Clarke generalized directional derivative $f^\circ(\boldsymbol{x}; \boldsymbol{d})$ always exists for a locally Lipschitz continuous function f. If f is convex $\partial f(\boldsymbol{x})$ coincides with the classical subdifferential of convex function (cf. [15]), in other words the set of $\boldsymbol{\xi} \in \mathbb{R}^n$ satisfying

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \quad \text{for all } \boldsymbol{y} \in \mathbb{R}^n.$$

Furthermore, if f is smooth $\partial f(\boldsymbol{x})$ reduces to $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$. Smoothness is critical here as the function

$$g(x) = \begin{cases} 0, & x = 0 \\ x^2 \cos(\frac{1}{x}), & x \neq 0 \end{cases} \tag{1}$$

shows. Function g is locally Lipschitz continuous and differentiable everywhere but nonsmooth (not continuously differentiable) and $\partial g(0) \neq \{\nabla g(0)\}$ (see appendix A).

The following properties derived in [4] are characteristic to the generalized directional derivative and subdifferential.

THEOREM 2.2. *If $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz continuous at $\boldsymbol{x} \in \mathbb{R}^n$, then*

(i) $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ *is positively homogeneous, subadditive and Lipschitz continuous function such that* $f^\circ(\boldsymbol{x}; -\boldsymbol{d}) = (-f)^\circ(\boldsymbol{x}; \boldsymbol{d})$.

(ii) $\partial f(\boldsymbol{x})$ *is a nonempty, convex and compact set.*

(iii) $f^\circ(\boldsymbol{x}; \boldsymbol{d}) = \max\{\boldsymbol{\xi}^T \boldsymbol{d} \mid \boldsymbol{\xi} \in \partial f(\boldsymbol{x})\}$ *for all* $\boldsymbol{d} \in \mathbb{R}^n$.

The subdifferential can be constructed as a convex hull of all possible limits of gradients at point $\boldsymbol{x}_i$ converging to $\boldsymbol{x}$. Let

$$\Omega_f = \{\boldsymbol{x} \in \mathbb{R}^n \mid f \text{ is not differentiable at the point } \boldsymbol{x}\}$$

be the set of points where f is not differentiable. By Rademacher's Theorem [10] a function which is Lipschitz continuous on a set $U \subseteq \mathbb{R}^n$ is differentiable almost everywhere on U, in other words, meas(Ω_f)=0.

THEOREM 2.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x} \in \mathbb{R}^n$. Then*

$$\partial f(\boldsymbol{x}) = \operatorname{conv} \{\boldsymbol{\xi} \in \mathbb{R}^n \mid \exists(\boldsymbol{x}_i) \subset \mathbb{R}^n \backslash \Omega_f \textit{ s.t. } \boldsymbol{x}_i \to \boldsymbol{x} \textit{ and } \nabla f(\boldsymbol{x}_i) \to \boldsymbol{\xi}\},$$

where conv *denotes the convex hull of a set.*

PROOF. See, for example, [10, pp. 50–51]. □

In order to maintain equalities instead of inclusions in subderivation rules we need the following regularity property.

DEFINITION 2.4. The function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *subdifferentially regular* at $\boldsymbol{x} \in \mathbb{R}^n$ if it is locally Lipschitz continuous at $\boldsymbol{x}$ and for all $\boldsymbol{d} \in \mathbb{R}^n$ the classical directional derivative

$$f'(\boldsymbol{x}; \boldsymbol{d}) = \lim_{t \downarrow 0} \frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t}$$

exists and $f'(\boldsymbol{x}; \boldsymbol{d}) = f^\circ(\boldsymbol{x}; \boldsymbol{d})$.

Note, that the equality $f'(\boldsymbol{x}; \boldsymbol{d}) = f^\circ(\boldsymbol{x}; \boldsymbol{d})$ is not necessarily valid in general even if $f'(\boldsymbol{x}; \boldsymbol{d})$ exists. This is the case, for instance, with concave nonsmooth functions. For example, the function $f(x) = -|x|$ has the directional derivative $f'(0; 1) = -1$, but the generalized directional derivative is $f^\circ(0; 1) = 1$. However, convexity, as well as smoothness implies subdifferential regularity [4]. Furthermore, it is easy to show that a necessary and sufficient condition for convexity is that for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we have

$$\begin{aligned} f(\boldsymbol{y}) - f(\boldsymbol{x}) &\geq f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \\ &= f'(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}). \end{aligned} \tag{2}$$

Next we present two subderivation rules of composite functions, namely the finite maximum and positive linear combination of subdifferentially regular functions.

THEOREM 2.5. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x}$ for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) := \max \{f_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}$$

is locally Lipschitz continuous at $\boldsymbol{x}$ and

$$\partial f(\boldsymbol{x}) \subseteq \operatorname{conv} \{\partial f_i(\boldsymbol{x}) \mid f_i(\boldsymbol{x}) = f(\boldsymbol{x}),\ i = 1, \ldots, m\} \tag{3}$$

In addition, if f_i is subdifferentially regular at $\boldsymbol{x}$ for all $i = 1, \ldots, m$, then f is also subdifferentially regular at $\boldsymbol{x}$ and equality holds in (3).

PROOF. See, for example, [4, p. 47]. □

THEOREM 2.6. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous at $\boldsymbol{x}$ and $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) := \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x})$$

is locally Lipschitz continuous at $\boldsymbol{x}$ *and*

$$\partial f(\boldsymbol{x}) \subseteq \sum_{i=1}^{m} \lambda_i \partial f_i(\boldsymbol{x}). \tag{4}$$

In addition, if f_i *is subdifferentially regular at* $\boldsymbol{x}$ *and* $\lambda_i \geq 0$ *for all* $i = 1, \ldots, m$*, then* f *is also subdifferentially regular at* $\boldsymbol{x}$ *and equality holds in (4).*

PROOF. See, for example, [4, pp. 39–40]. □

The following two results generalize the classical Mean-Value Theorem and the Chain Rule, respectively.

THEOREM 2.7. *Let* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *be such that* $\boldsymbol{x} \neq \boldsymbol{y}$ *and let the function* f *be locally Lipschitz continuous on an open set* $U \subseteq \mathbb{R}^n$ *such that the line segment* $[\boldsymbol{x}, \boldsymbol{y}] \subset U$*. Then there exists a point* $\boldsymbol{z} \in (\boldsymbol{x}, \boldsymbol{y})$ *such that*

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \in \partial f(\boldsymbol{z})^T (\boldsymbol{y} - \boldsymbol{x}).$$

PROOF. See, for example, [4, pp. 41–42]. □

THEOREM 2.8. *Let* $h : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous at* $\boldsymbol{x}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *be locally Lipschitz continuous at* $h(\boldsymbol{x})$*. Then the composite function* $f = g \circ h : \mathbb{R}^n \to \mathbb{R}$ *is also locally Lipschitz continuous at* $\boldsymbol{x}$ *and one has*

$$\partial f(\boldsymbol{x}) \subseteq \operatorname{conv} \{\partial g(h(\boldsymbol{x})) \partial h(\boldsymbol{x})\}.$$

PROOF. See, for example, [4, pp. 72–73]. □

Next we shall give the basic unconstrained optimality condition.

THEOREM 2.9. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous at* $\boldsymbol{x}^*$*. If* f *attains its local minimum at* $\boldsymbol{x}^*$*, then*

$$\mathbf{0} \in \partial f(\boldsymbol{x}^*).$$

If, in addition, f *is convex, then the above condition is also sufficient for* $\boldsymbol{x}^*$ *to be a global minimum.*

PROOF. See, for example, [10, pp. 70–71]. □

Now we shall present a theorem and two lemmas that are used later.

THEOREM 2.10. *Let* $\varepsilon > 0$*, function* $f : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz continuous at* $\boldsymbol{x}$ *and* $\boldsymbol{d} \in \mathbb{R}^n$*,* $\boldsymbol{d} \neq \mathbf{0}$*. Then*

$$f^\circ(\boldsymbol{x}; \boldsymbol{d}) - \varepsilon \leq \limsup \left\{ \nabla f(\boldsymbol{y})^T \boldsymbol{d} \mid \boldsymbol{y} \to \boldsymbol{x},\, \boldsymbol{y} \notin \Omega_f \right\}.$$

PROOF. See, for example, [10, pp. 51–52]. □

LEMMA 2.11. *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a point, where $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz continuous and differentiable. Let K be the Lipschitz constant of the function f at the point $\boldsymbol{x}$. Then the function $\boldsymbol{d} \mapsto f'(\boldsymbol{x}; \boldsymbol{d})$ is positively homogeneous and Lipschitz continuous with the constant K.*

PROOF. Since f is differentiable at the point $\boldsymbol{x}$ the directional derivatives $f'(\boldsymbol{x}; \boldsymbol{d})$ exist for all $\boldsymbol{d} \in \mathbb{R}^n$. Let $\lambda > 0$, then

$$\begin{aligned} f'(\boldsymbol{x}; \lambda \boldsymbol{d}) &= \lim_{t \downarrow 0} \frac{f(\boldsymbol{x} + \lambda \boldsymbol{d} t) - f(\boldsymbol{x})}{t} = \lim_{t \downarrow 0} \lambda \frac{f(\boldsymbol{x} + \lambda \boldsymbol{d} t) - f(\boldsymbol{x})}{\lambda t} \\ &= \lambda \lim_{t \downarrow 0} \frac{f(\boldsymbol{x} + \lambda \boldsymbol{d} t) - f(\boldsymbol{x})}{\lambda t} = \lambda f'(\boldsymbol{x}; \boldsymbol{d}), \end{aligned}$$

which proves the positive homogeneity.

Let $\boldsymbol{u}, \boldsymbol{w} \in \mathbb{R}^n$ be arbitrary. Since f is locally Lipschitz continuous there exists $\varepsilon > 0$ such that the Lipschitz condition holds in $B(\boldsymbol{x}; \varepsilon)$. Furthermore, there exists $t^0 > 0$ such that $\boldsymbol{x} + \boldsymbol{w}t, \boldsymbol{x} + \boldsymbol{u}t \in B(\boldsymbol{x}; \varepsilon)$ when $0 < t < t^0$. Then

$$f(\boldsymbol{x} + \boldsymbol{u}t) - f(\boldsymbol{x} + \boldsymbol{w}t) \leq Kt \left\| \boldsymbol{u} - \boldsymbol{w} \right\|,$$

and, thus,

$$\lim_{t \downarrow 0} \frac{f(\boldsymbol{x} + \boldsymbol{u}t) - f(\boldsymbol{x})}{t} \leq \lim_{t \downarrow 0} \frac{f(\boldsymbol{x} + \boldsymbol{w}t) - f(\boldsymbol{x})}{t} + K \left\| \boldsymbol{u} - \boldsymbol{w} \right\|$$

whence

$$f'(\boldsymbol{x}; \boldsymbol{u}) - f'(\boldsymbol{x}; \boldsymbol{w}) \leq K \left\| \boldsymbol{u} - \boldsymbol{w} \right\|.$$

Reversing the roles of $\boldsymbol{u}$ and $\boldsymbol{w}$ we obtain

$$f'(\boldsymbol{x}; \boldsymbol{w}) - f'(\boldsymbol{x}; \boldsymbol{u}) \leq K \left\| \boldsymbol{u} - \boldsymbol{w} \right\|.$$

Thus

$$|f'(\boldsymbol{x}; \boldsymbol{w}) - f'(\boldsymbol{x}; \boldsymbol{u})| \leq K \left\| \boldsymbol{u} - \boldsymbol{w} \right\|$$

completing the proof of the Lipschitz continuity. $\square$

The *level set* of f with a parameter $\alpha \in \mathbb{R}$ is defined as

$$\operatorname{lev}_\alpha f := \{ \boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) \leq \alpha \}$$

LEMMA 2.12. *Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and on $[\boldsymbol{x}, \boldsymbol{y}]$ locally Lipschitz continuous function f be such that $f(\boldsymbol{x}) < f(\boldsymbol{y})$. Then, there exists a point $\bar{\boldsymbol{x}} = \lambda \boldsymbol{x} + (1 - \lambda) \boldsymbol{y}$, $\lambda \in (0, 1)$ such that $f(\bar{\boldsymbol{x}}) > f(\boldsymbol{x})$ and $f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x}) > 0$.*

PROOF. Consider the nonempty set $A = \operatorname{lev}_{f(\boldsymbol{x})} f \cap [\boldsymbol{x}, \boldsymbol{y}]$. Since level sets of a continuous function are closed sets and $[\boldsymbol{x}, \boldsymbol{y}]$ is compact, the set A is a compact set. Since function $g(\boldsymbol{w}) := \|\boldsymbol{w} - \boldsymbol{y}\|$ is continuous, it has a minimum on the set A according to the well-known Weierstrass Theorem. Let this minimum point be $\boldsymbol{z}$. Then $\boldsymbol{z}$ is the nearest point to $\boldsymbol{y}$ on the set A and the continuity of function f implies $f(\boldsymbol{z}) = f(\boldsymbol{x})$. Also, $\boldsymbol{z} \neq \boldsymbol{y}$ since $f(\boldsymbol{x}) < f(\boldsymbol{y})$. The Mean-Value Theorem implies that there exist $\bar{\boldsymbol{z}} \in (\boldsymbol{z}, \boldsymbol{y})$ and $\boldsymbol{\xi} \in \partial f(\bar{\boldsymbol{z}})$ such that

$$f(\boldsymbol{y}) - f(\boldsymbol{z}) = \boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{z}).$$

Since $f(\boldsymbol{z}) < f(\boldsymbol{y})$ we have

$$0 < f(\boldsymbol{y}) - f(\boldsymbol{z}) = \boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{z}) \leq f^\circ(\bar{\boldsymbol{z}}; \boldsymbol{y} - \boldsymbol{z}) \leq f^\circ(\bar{\boldsymbol{z}}; \boldsymbol{y} - \boldsymbol{x}),$$

where the last inequality follows from positive homogeneity and inequality $\|\boldsymbol{y} - \boldsymbol{z}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\|$. By the choice of the point $\boldsymbol{z}$ we know that $f(\boldsymbol{z}) < f(\bar{\boldsymbol{z}})$ since $\bar{\boldsymbol{z}} \in (\boldsymbol{z}, \boldsymbol{y})$. Choosing $\bar{\boldsymbol{x}} = \bar{\boldsymbol{z}}$ we have $f(\boldsymbol{x}) = f(\boldsymbol{z}) < f(\bar{\boldsymbol{x}})$ and the lemma has been proven. □

3 Generalized Pseudoconvexity

The most famous definition of pseudoconvexity for smooth functions was introduced in [11].

DEFINITION 3.1. A continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is *pseudoconvex*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) < f(\boldsymbol{x}) \qquad \text{implies} \qquad \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) < 0.$$

The main result for a smooth pseudoconvex function f is that the convexity assumption of Theorem 2.9 can be weakened, in other words, a smooth pseudoconvex function f attains a global minimum at $\boldsymbol{x}^*$, if and only if $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$ (see [11]).

Lately, the concept of pseudoconvexity has been extended for nonsmooth case by many authors (see e.g., [1, 14] and the references therein). One way to do this is the usage of directional derivatives. The Dini directional derivatives were used, for example, by Diewert [5], Komlósi [8] and Borde and Crouzeix [3]. In [9] this idea was generalized for lower semicontinuous functions via h-pseudoconvexity, where $h(\boldsymbol{x}, \boldsymbol{d})$ is any real-valued bifunction, that is, for example, any directional derivative. In this paper we use the definition by Hiriart-Urruty [7] for locally Lipschitz continuous functions.

DEFINITION 3.2. A function $f : \mathbb{R}^n \to \mathbb{R}$ is f°-*pseudoconvex*, if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) < f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) < 0.$$

Note that due to (2) a convex function is always f°-pseudoconvex. The next result shows that f°-pseudoconvexity is a natural extension of pseudoconvexity.

THEOREM 3.3. *If f is smooth, then f is f°-pseudoconvex, if and only if f is pseudoconvex.*

PROOF. Follows immediately from Theorem 2.2 (iii), since for a smooth f we have $f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) = f'(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) = \nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x})$. □

Sometimes the reasoning chain in the definition of f°-pseudoconvexity needs to be converted.

LEMMA 3.4. *A locally Lipschitz continuous function f is f°-pseudoconvex, if and only if for all $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$*

$$f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) \geq 0 \qquad \textit{implies} \qquad f(\boldsymbol{y}) \geq f(\boldsymbol{x}).$$

PROOF. Follows directly from the definition of f°-pseudoconvexity. □

The important sufficient extremum property of pseudoconvexity remains also for f°-pseudoconvexity.

THEOREM 3.5. *An f°-pseudoconvex f attains its global minimum at $\boldsymbol{x}^*$, if and only if*

$$\boldsymbol{0} \in \partial f(\boldsymbol{x}^*).$$

PROOF. If f attains its global minimum at $\boldsymbol{x}^*$, then by Theorem 2.9 we have $\boldsymbol{0} \in \partial f(\boldsymbol{x}^*)$. On the other hand, if $\boldsymbol{0} \in \partial f(\boldsymbol{x}^*)$ and $\boldsymbol{y} \in \mathbb{R}^n$, then by Definition 2.1

$$f^\circ(\boldsymbol{x}^*;\boldsymbol{y}-\boldsymbol{x}^*) \geq \boldsymbol{0}^T(\boldsymbol{y}-\boldsymbol{x}^*) = 0$$

and, thus by Lemma 3.4 we have

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}^*).$$

□

The following example shows, that f°-pseudoconvexity is a more general property than pseudoconvexity.

EXAMPLE 3.1. Define $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) := \min\{|x|, x^2\}$. Then f is clearly locally Lipschitz continuous but not convex nor pseudoconvex. However, for all $y > x$ we have

$$f^\circ(x;y-x) = \begin{cases} -1, & x \in (-\infty,-1] \\ 2x, & x \in (-1,1] \\ 1, & x \in (1,\infty), \end{cases}$$

and thus, due to the symmetricity of the function f and Lemma 3.4, f is f°-pseudoconvex. Furthermore for the unique global minimum $x^* = 0$ we have $\partial f(x^*) = \{0\}$.

The notion of monotonicity is closely related to convexity.

DEFINITION 3.6. The generalized directional derivative f° is called *pseudomonotone*, if for all $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$

$$f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) \geq 0 \qquad \text{implies} \qquad f^\circ(\boldsymbol{y};\boldsymbol{x}-\boldsymbol{y}) \leq 0$$

or, equivalently

$$f^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) > 0 \qquad \text{implies} \qquad f^\circ(\boldsymbol{y}; \boldsymbol{x}-\boldsymbol{y}) < 0.$$

Furthermore, f° is *strictly pseudomonotone*, if

$$f^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) \geq 0 \qquad \text{implies} \qquad f^\circ(\boldsymbol{y}; \boldsymbol{x}-\boldsymbol{y}) < 0.$$

THEOREM 3.7. *If f is locally Lipschitz continuous such that f° is pseudomonotone, then f is f°-pseudoconvex.*

PROOF. Let us, on the contrary, assume that f is not f°-pseudoconvex. Then there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $f(\boldsymbol{y}) < f(\boldsymbol{x})$ and

$$f^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) \geq 0. \tag{5}$$

Then by the Mean-Value Theorem 2.7 there exists $\hat{\lambda} \in (0,1)$ such that $\hat{\boldsymbol{x}} = \boldsymbol{x} + \hat{\lambda}(\boldsymbol{y}-\boldsymbol{x})$ and

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) \in \partial f(\hat{\boldsymbol{x}})^T(\boldsymbol{x}-\boldsymbol{y}).$$

This means that due to the definition of the Clarke subdifferential there exists $\hat{\boldsymbol{\xi}} \in \partial f(\hat{\boldsymbol{x}})$ such that

$$0 < f(\boldsymbol{x}) - f(\boldsymbol{y}) = \hat{\boldsymbol{\xi}}^T(\boldsymbol{x}-\boldsymbol{y}) \leq f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x}-\boldsymbol{y}). \tag{6}$$

On the other hand, from (5) and the positive homogeneity of $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ (see Theorem 2.2 (i)) we deduce that

$$f^\circ(\boldsymbol{x}; \hat{\boldsymbol{x}}-\boldsymbol{x}) = f^\circ(\boldsymbol{x}; \hat{\lambda}(\boldsymbol{y}-\boldsymbol{x})) = \hat{\lambda} f^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) \geq 0.$$

Then the pseudomonotonicity, the positive homogeneity of $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ and (6) imply that

$$0 \geq f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x}-\hat{\boldsymbol{x}}) = \hat{\lambda} f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x}-\boldsymbol{y}) > 0,$$

which is impossible. Thus f is f°-pseudoconvex. □

The converse is true too. Few lemmas is needed before the proof.

LEMMA 3.8. *Let f be an f°-pseudoconvex function, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\bar{\lambda} \in (0,1)$. Denote $\bar{\boldsymbol{x}} = \bar{\lambda}\boldsymbol{x} + (1-\bar{\lambda})\boldsymbol{y}$. Then, $f(\bar{\boldsymbol{x}}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$*

PROOF. On the contrary assume that $f(\bar{\boldsymbol{x}}) > \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$. Since f is f°-pseudoconvex and $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ is positively homogeneous by Theorem 2.2 (i), we have

$$0 > f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{x}-\bar{\boldsymbol{x}}) = f^\circ(\bar{\boldsymbol{x}}; (1-\bar{\lambda})(\boldsymbol{x}-\boldsymbol{y})) = (1-\bar{\lambda}) f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{x}-\boldsymbol{y})$$

and thus

$$f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{x}-\boldsymbol{y}) < 0.$$

Correspondingly, we obtain

$$0 > f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{y} - \bar{\boldsymbol{x}}) = f^\circ(\bar{\boldsymbol{x}}; \bar{\lambda}(\boldsymbol{y} - \boldsymbol{x})) = \bar{\lambda} f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x})$$

and thus

$$f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x}) < 0.$$

Since $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ is subadditive by Theorem 2.2 (i), we have

$$0 > f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{x} - \boldsymbol{y}) + f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x}) \geq f^\circ(\bar{\boldsymbol{x}}; (\boldsymbol{x} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{x})) = f^\circ(\bar{\boldsymbol{x}}; \boldsymbol{0}) = 0,$$

which is impossible. In other words, $f(\bar{\boldsymbol{x}}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$. □

LEMMA 3.9. *Let f be an f°-pseudoconvex function. Then there exist no points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, which satisfy the following conditions*

(i) $f(\boldsymbol{x}) = f(\boldsymbol{y})$ *and*

(ii) $f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0$.

PROOF. On the contrary, assume that there exist points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\delta > 0$ such that $f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) = \delta$ and $f(\boldsymbol{x}) = f(\boldsymbol{y})$. Since f is locally Lipschitz continuous there exist $\varepsilon, K > 0$ such that K is the Lipschitz constant in the ball $B(\boldsymbol{x}; \varepsilon)$. Since $f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) = \delta$ Theorem 2.10 implies that there exists a sequence $(\boldsymbol{z}^i)$ of points where f is differentiable and $I \in \mathbb{N}$ such that $\boldsymbol{z}^i \to \boldsymbol{x}$ and

$$f'(\boldsymbol{z}^i; \boldsymbol{y} - \boldsymbol{x}) = \nabla f(\boldsymbol{z}^i)^T(\boldsymbol{y} - \boldsymbol{x}) > \frac{\delta}{2} \tag{7}$$

holds when $i \geq I$. Let

$$\hat{\varepsilon} = \min\left\{\varepsilon, \frac{\delta}{2K}\right\}$$

and $\boldsymbol{z} \in B(\boldsymbol{x}; \hat{\varepsilon}) \cap \{(\boldsymbol{z}^i) \mid i \geq I\}$. According to Lemma 2.11 $f'(\boldsymbol{z}; \cdot)$ is Lipschitz continuous with the constant K. Hence,

$$\begin{aligned} |f'(\boldsymbol{z}; \boldsymbol{y} - \boldsymbol{x}) - f'(\boldsymbol{z}; \boldsymbol{y} - \boldsymbol{z})| &\leq K\|\boldsymbol{y} - \boldsymbol{x} - (\boldsymbol{y} - \boldsymbol{z})\| \\ &= K\|\boldsymbol{z} - \boldsymbol{x}\| < K\frac{\delta}{2K} = \frac{\delta}{2}. \end{aligned} \tag{8}$$

Thus, $f'(\boldsymbol{z}; \boldsymbol{y} - \boldsymbol{z}) > 0$ according to (7) and (8). Since $f'(\boldsymbol{z}; \boldsymbol{y} - \boldsymbol{z}) > 0$ there exists $\mu \in (0, 1)$ such that

$$f(\mu\boldsymbol{z} + (1 - \mu)\boldsymbol{y}) > f(\boldsymbol{z}). \tag{9}$$

Since $f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) = \delta$, Theorem 2.2 (i) implies that there exists $\bar{\varepsilon} > 0$ such that $f^\circ(\boldsymbol{x}; \boldsymbol{d}) > 0$ when $\boldsymbol{d} \in B(\boldsymbol{y} - \boldsymbol{x}; \bar{\varepsilon})$. Let $\bar{\boldsymbol{z}} \in B(\boldsymbol{y}; \bar{\varepsilon})$. Since

$$\|\bar{\boldsymbol{z}} - \boldsymbol{x} - (\boldsymbol{y} - \boldsymbol{x})\| = \|\bar{\boldsymbol{z}} - \boldsymbol{y}\| < \bar{\varepsilon},$$

it follows that $\bar{\boldsymbol{z}}-\boldsymbol{x} \in B(\boldsymbol{y}-\boldsymbol{x};\bar{\varepsilon})$. Thus, $f^\circ(\boldsymbol{x};\bar{\boldsymbol{z}}-\boldsymbol{x}) > 0$ and the f°-pseudoconvexity of the function f implies $f(\bar{\boldsymbol{z}}) \geq f(\boldsymbol{x}) = f(\boldsymbol{y})$. Thus, $\boldsymbol{y}$ is a local minimum for the function f and Theorem 2.9 implies that $\mathbf{0} \in \partial f(\boldsymbol{y})$. Due to Theorem 3.5 $\boldsymbol{y}$ is also a global minimum. Thus, we have $f(\boldsymbol{y}) \leq f(\boldsymbol{z})$ and the inequality (9) implies that

$$f(\mu\boldsymbol{z} + (1-\mu)\boldsymbol{y}) > \max\{f(\boldsymbol{z}), f(\boldsymbol{y})\},$$

which is impossible by Lemma 3.8. □

REMARK 3.1. It is good to note that differentiability at the point $\boldsymbol{x}$ is crucial in Lemma 2.11. This allows us to assume that directional derivatives $f'(\boldsymbol{x};\boldsymbol{d})$ exist at $\boldsymbol{x}$ which was needed in Lemma 3.9. Unlike the convexity, the f°-pseudoconvexity does not guarantee that directional derivatives exist at every point. An example of f°-pseudoconvex function for which directional derivatives do not exist at every point is presented in Appendix B.

Now we are ready to prove the converse result of Theorem 3.7.

THEOREM 3.10. *The generalized directional derivative of a f°-pseudoconvex function is pseudomonotone.*

PROOF. Let f be f°-pseudoconvex and, on the contrary, assume that there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) \geq 0$ and $f^\circ(\boldsymbol{y};\boldsymbol{x}-\boldsymbol{y}) > 0$. Then, by f°-pseudoconvexity $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ and $f(\boldsymbol{y}) \leq f(\boldsymbol{x})$, hence $f(\boldsymbol{x}) = f(\boldsymbol{y})$. Thus, we have $f^\circ(\boldsymbol{y};\boldsymbol{x}-\boldsymbol{y}) > 0$ and $f(\boldsymbol{x}) = f(\boldsymbol{y})$, which contradicts Lemma 3.9. □

In what follows we consider how to verify the f°-pseudoconvexity in practice. Before that, however, we need the following result.

LEMMA 3.11. *A locally Lipschitz continuous function $g : \mathbb{R} \to \mathbb{R}$ is f°-pseudoconvex and strictly increasing, if and only if $\varsigma > 0$ for all $\varsigma \in \partial g(x)$ and $x \in \mathbb{R}$.*

PROOF. Suppose first that g is both f°-pseudoconvex and strictly increasing and let $v < 0$. Then for every $x \in \mathbb{R}$ we have $g(x+v) < g(x)$ and due to f°-pseudoconvexity $g^\circ(x;v) < 0$. By the definition of the subdifferential for all $\varsigma \in \partial g(x)$ we have

$$\varsigma v \leq g^\circ(x;v) < 0,$$

which implies $\varsigma > 0$.

On the other hand, let all the subgradients of g be positive. We first prove that g is strictly increasing. Suppose, on the contrary, that there exist $y, x \in \mathbb{R}$ such that $y < x$ and $g(y) \geq g(x)$. By the Mean-Value Theorem 2.7 there exists $\hat{x} \in (y, x)$ such that

$$g(x) - g(y) \in \partial g(\hat{x})(x-y).$$

This means that there exists $\hat{\varsigma} \in \partial g(\hat{x})$ such that $\hat{\varsigma} > 0$ and

$$0 \geq g(x) - g(y) = \hat{\varsigma}(x-y) > 0,$$

which is impossible. Thus, g is strictly increasing.

Since g is strictly increasing we have $g(y) < g(x)$ if and only if $y < x$, where $x, y \in \mathbb{R}$. Thus, to prove f°-pseudoconvexity we need to show that $y < x$ implies $f^\circ(x; y - x) < 0$. Let $x, y \in \mathbb{R}$ be arbitrary such that $y < x$. By Theorem 2.2 (iii)

$$f^\circ(x; y - x) = \max\{\varsigma(y - x) \mid \varsigma \in \partial f(x)\} < 0$$

which proves the f°-pseudoconvexity. □

THEOREM 3.12. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be f°-pseudoconvex and $g : \mathbb{R} \to \mathbb{R}$ be f°-pseudoconvex and strictly increasing. Then the composite function $f := g \circ h : \mathbb{R}^n \to \mathbb{R}$ is also f°-pseudoconvex.*

PROOF. According to Theorem 2.8 function f is locally Lipschitz continuous. Suppose now that $f(\boldsymbol{y}) < f(\boldsymbol{x})$. Then $g(h(\boldsymbol{y})) = f(\boldsymbol{y}) < f(\boldsymbol{x}) = g(h(\boldsymbol{x}))$ and since g is strictly increasing we have

$$h(\boldsymbol{y}) < h(\boldsymbol{x}). \tag{10}$$

From Theorems 2.2 (iii) and 2.8 we deduce that

$$\begin{aligned} f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) &= \max\{\boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \mid \boldsymbol{\xi} \in \partial f(\boldsymbol{x})\} \\ &\leq \max\left\{\boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \mid \boldsymbol{\xi} \in \operatorname{conv}\{\partial g(h(\boldsymbol{x}))\partial h(\boldsymbol{x})\}\right\} \end{aligned} \tag{11}$$

Due to the definition of a convex hull the right hand side of (11) is equivalent to

$$\begin{aligned} &\max\Big\{\Big(\sum_{i=1}^m \lambda_i \varsigma_i \boldsymbol{\zeta}_i\Big)^T (\boldsymbol{y} - \boldsymbol{x}) \mid \varsigma_i \in \partial g(h(\boldsymbol{x})),\ \boldsymbol{\zeta}_i \in \partial h(\boldsymbol{x}),\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1\Big\} \\ &\leq \max\Big\{\Big(\sum_{i=1}^m \lambda_i \varsigma_i\Big) \cdot \max_{\boldsymbol{\zeta}_i \in \partial h(\boldsymbol{x})} \boldsymbol{\zeta}_i^T (\boldsymbol{y} - \boldsymbol{x}) \mid \varsigma_i \in \partial g(h(\boldsymbol{x})),\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1\Big\} \\ &= \max\Big\{\Big(\sum_{i=1}^m \lambda_i \varsigma_i\Big) h^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \mid \varsigma_i \in \partial g(h(\boldsymbol{x})),\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1\Big\}, \end{aligned}$$

since by Lemma 3.11 we have $\varsigma_i > 0$ for all $i = 1, \ldots, m$ and thus

$$\sum_{i=1}^m \lambda_i \varsigma_i > 0.$$

On the other hand, since h is f°-pseudoconvex, (10) implies that $h^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) < 0$. Then

$$\begin{aligned} f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) &\leq \max\Big\{\Big(\sum_{i=1}^m \lambda_i \varsigma_i\Big) h^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \mid \varsigma_i \in \partial g(h(\boldsymbol{x})),\ \lambda_i \geq 0,\ \sum_{i=1}^m \lambda_i = 1\Big\} \\ &< 0 \end{aligned}$$

thus f is f°-pseudoconvex. □

THEOREM 3.13. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be f°-pseudoconvex for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) := \max\{f_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}$$

is also f°-pseudoconvex.

PROOF. According to Theorem 2.5 f is locally Lipschitz continuous. Suppose that $f(\boldsymbol{y}) < f(\boldsymbol{x})$. Define the index set

$$I(\boldsymbol{x}) := \{i \in \{1, \ldots, m\} \mid f_i(\boldsymbol{x}) = f(\boldsymbol{x})\}.$$

Then for all $i \in I(\boldsymbol{x})$ we have

$$f_i(\boldsymbol{y}) \leq f(\boldsymbol{y}) < f(\boldsymbol{x}) = f_i(\boldsymbol{x}). \tag{12}$$

From Theorem 2.2 (iii) and 2.5, the definition of a convex hull, f°-pseudoconvexity of f_i, and (12) we deduce that

$$\begin{aligned}
f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) &= \max\{\boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \mid \boldsymbol{\xi} \in \partial f(\boldsymbol{x})\} \\
&\leq \max\left\{\boldsymbol{\xi}^T(\boldsymbol{y} - \boldsymbol{x}) \mid \boldsymbol{\xi} \in \operatorname{conv}\{\partial f_i(\boldsymbol{x})) \mid i \in I(\boldsymbol{x})\}\right\} \\
&= \max\left\{\Big(\sum_{i \in I(\boldsymbol{x})} \lambda_i \boldsymbol{\xi}_i\Big)^T (\boldsymbol{y} - \boldsymbol{x}) \mid \boldsymbol{\xi}_i \in \partial f(\boldsymbol{x}),\ \lambda_i \geq 0,\ \sum_{i \in I(\boldsymbol{x})} \lambda_i = 1\right\} \\
&\leq \max\left\{\sum_{i \in I(\boldsymbol{x})} \lambda_i \cdot \max_{\boldsymbol{\xi}_i \in \partial f_i(\boldsymbol{x})} \boldsymbol{\xi}_i^T(\boldsymbol{y} - \boldsymbol{x}) \mid \lambda_i \geq 0,\ \sum_{i \in I(\boldsymbol{x})} \lambda_i = 1\right\} \\
&= \max\left\{\sum_{i \in I(\boldsymbol{x})} \lambda_i f_i^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \mid \lambda_i \geq 0,\ \sum_{i \in I(\boldsymbol{x})} \lambda_i = 1\right\} < 0.
\end{aligned}$$

Thus, f is f°-pseudoconvex. □

Due to the fact that the sum of f°-pseudoconvex functions is not necessarily f°-pseudoconvex we need the following new property.

DEFINITION 3.14. The functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are said to be *additively strictly monotone*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda_i \geq 0$, $i = 1, \ldots, m$

$$\sum_{i=1}^m \lambda_i f_i(\boldsymbol{y}) < \sum_{i=1}^m \lambda_i f_i(\boldsymbol{x}) \qquad \text{implies} \qquad f_i(\boldsymbol{y}) < f_i(\boldsymbol{x}).$$

THEOREM 3.15. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be f°-pseudoconvex and additively strictly monotone, and let $\lambda_i \geq 0$ for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) := \sum_{i=1}^m \lambda_i f_i(\boldsymbol{x})$$

is f°-pseudoconvex.

PROOF. According to Theorem 2.6 f is locally Lipschitz continuous. Suppose that $f(\boldsymbol{y}) < f(\boldsymbol{x})$. Then the additive strict monotonicity implies that for all $i = 1, \ldots, m$ we have

$$f_i(\boldsymbol{y}) < f_i(\boldsymbol{x}). \tag{13}$$

From Theorem 2.2 (iii) and 2.6, nonnegativity of λ_i, f°-pseudoconvexity of f_i, and (13) we deduce that

$$\begin{aligned} f^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) &= \max\{\boldsymbol{\xi}^T(\boldsymbol{y}-\boldsymbol{x}) \mid \boldsymbol{\xi} \in \partial f(\boldsymbol{x})\} \\ &\leq \max\{\boldsymbol{\xi}^T(\boldsymbol{y}-\boldsymbol{x}) \mid \boldsymbol{\xi} \in \sum_{i=1}^{m} \lambda_i \partial f_i(\boldsymbol{x})\} \\ &= \max\left\{\Big(\sum_{i=1}^{m} \lambda_i \boldsymbol{\xi}_i\Big)^T (\boldsymbol{y}-\boldsymbol{x}) \mid \boldsymbol{\xi}_i \in \partial f_i(\boldsymbol{x})\right\} \\ &\leq \sum_{i=1}^{m} \lambda_i \cdot \max_{\boldsymbol{\xi}_i \in \partial f_i(\boldsymbol{x})} \boldsymbol{\xi}_i^T(\boldsymbol{y}-\boldsymbol{x}) \\ &= \sum_{i=1}^{m} \lambda_i f_i^\circ(\boldsymbol{x}; \boldsymbol{y}-\boldsymbol{x}) < 0. \end{aligned}$$

Thus, f is f°-pseudoconvex. □

4 Generalized Quasiconvexity

The notion of quasiconvexity is the most widely used generalization of convexity, and, thus, there exist various equivalent definitions and characterizations. Next we recall the most commonly used definition of quasiconvexity (see e.g. [1]).

DEFINITION 4.1. The function $f : \mathbb{R}^n \to \mathbb{R}$ is *quasiconvex*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.$$

REMARK 4.1. Lemma 3.8 implies that f°-pseudoconvex function is also quasiconvex.

Note, that unlike pseudoconvexity, the previous definition of quasiconvexity does not require differentiability. Next we will give a well-known important geometrical characterization to quasiconvexity.

THEOREM 4.2. *A function f is quasiconvex, if and only if the level set $\mathrm{lev}_\alpha f$ is a convex set for all $\alpha \in \mathbb{R}$.*

PROOF. Let f be quasiconvex, $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{lev}_\alpha f$, $\lambda \in [0, 1]$ and $\alpha \in \mathbb{R}$. Then

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\} \leq \max\{\alpha, \alpha\} = \alpha,$$

thus $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathrm{lev}_\alpha f$.

On the other hand, let $\mathrm{lev}_\alpha f$ be a convex set for all $\alpha \in \mathbb{R}$. By choosing $\beta := \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$ we have $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{lev}_\beta f$. The convexity of $\mathrm{lev}_\beta f$ implies, that $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in \mathrm{lev}_\beta f$ for all $\lambda \in [0,1]$, in other words

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \le \beta = \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.$$

□

We give also a useful result concerning a finite maximum of quasiconvex functions.

THEOREM 4.3. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be quasiconvex at $\boldsymbol{x}$ for all $i = 1, \dots, m$. Then the function*

$$f(\boldsymbol{x}) := \max\{f_i(\boldsymbol{x}) \mid i = 1, \dots, m\}$$

is also quasiconvex.

PROOF. Follows directly from the definition of quasiconvexity. □

Also the concept of quasiconvexity has been studied by many authors (see [14] and references therein). The Dini directional derivatives were used in the characterization of quasiconvexity for radially lower semicontinuous functions in [5]. Analogously to the Definition 3.2 we can define the corresponding generalized concept, which is a special case of h-quasiconvexity defined by Komlósi [9] when h is the Clarke generalized directional derivative.

DEFINITION 4.4. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *f°-quasiconvex*, if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) \le f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \le 0.$$

Similarly to f°-pseudoconvexity, the reasoning chain may be converted.

LEMMA 4.5. *The locally Lipschitz continuous function f is f°-quasiconvex, if and only if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$*

$$f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0 \qquad \textit{implies} \qquad f(\boldsymbol{y}) > f(\boldsymbol{x}).$$

PROOF. Follows directly from the definition of f°-quasiconvexity. □

There is a way, similar to Definition 4.4, to express locally Lipschitz continuous and quasiconvex function.

DEFINITION 4.6. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *l-quasiconvex*, if it is locally Lipschitz continuous and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(\boldsymbol{y}) < f(\boldsymbol{x}) \qquad \text{implies} \qquad f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) \le 0.$$

REMARK 4.2. Definitions 4.4 and 4.6 imply that an f°-quasiconvex function is l-quasiconvex.

Next, we prove that l-quasiconvexity coincides with quasiconvexity in locally Lipschitz continuous case. This result can be found in [6].

THEOREM 4.7. *If a locally Lipschitz continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex then it is l-quasiconvex.*

PROOF. Let f be locally Lipschitz continuous and quasiconvex. Let $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$ be such that $f(\boldsymbol{z}) < f(\boldsymbol{x})$. Since local Lipschitz continuity implies continuity there exists $\varepsilon > 0$ such that $f(\boldsymbol{z}+\boldsymbol{d}) < f(\boldsymbol{x}+\boldsymbol{d})$ for all $\boldsymbol{d} \in B(\boldsymbol{0}; \varepsilon)$. For generalized directional derivative $f^\circ(\boldsymbol{x}; \boldsymbol{z} - \boldsymbol{x})$ we have

$$\begin{aligned} f^\circ(\boldsymbol{x}; \boldsymbol{z} - \boldsymbol{x}) &= \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \downarrow 0}} \frac{f(\boldsymbol{y} + t(\boldsymbol{z} - \boldsymbol{x})) - f(\boldsymbol{y})}{t} \\ &= \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \downarrow 0}} \frac{f(\boldsymbol{y} + t(\boldsymbol{z} - \boldsymbol{x} + \boldsymbol{y} - \boldsymbol{y})) - f(\boldsymbol{y})}{t} \\ &= \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \downarrow 0}} \frac{f((1-t)\boldsymbol{y} + t(\boldsymbol{z} + \boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{y})}{t} \end{aligned}$$

When $t \in (0, 1)$ and $\boldsymbol{y} - \boldsymbol{x} \in B(\boldsymbol{0}; \varepsilon)$ the quasiconvexity of f implies

$$\begin{aligned} & \frac{f((1-t)\boldsymbol{y} + t(\boldsymbol{z} + \boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{y})}{t} \\ \le \; & \frac{\max\{f(\boldsymbol{y}), f(\boldsymbol{z} + \boldsymbol{y} - \boldsymbol{x})\} - f(\boldsymbol{y} - \boldsymbol{x} + \boldsymbol{x})}{t} \\ = \; & \frac{\max\{0, f(\boldsymbol{z} + \boldsymbol{y} - \boldsymbol{x}) - f(\boldsymbol{x} + \boldsymbol{y} - \boldsymbol{x})\}}{t} = 0 \end{aligned}$$

Passing to the limit $t \to 0$ and $\boldsymbol{y} \to \boldsymbol{x}$ we get $f^\circ(\boldsymbol{x}; \boldsymbol{z} - \boldsymbol{x}) \le 0$. Thus, f is l-quasiconvex. □

THEOREM 4.8. *If function $f : \mathbb{R}^n \to \mathbb{R}$ is l-quasiconvex then it is quasiconvex.*

PROOF. On the contrary assume that an l-quasiconvex function f is not quasiconvex. Then there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\bar{\lambda} \in (0, 1)$ such that $f(\bar{\boldsymbol{x}}) > \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}$, where $\bar{\boldsymbol{x}} = \bar{\lambda}\boldsymbol{x} + (1 - \bar{\lambda})\boldsymbol{y}$. Without a loss of generality we may assume that $f(\boldsymbol{x}) \ge f(\boldsymbol{y})$. Lemma 2.12 implies that there exists $\tilde{\boldsymbol{x}} \in (\boldsymbol{x}, \bar{\boldsymbol{x}})$, for which

$$f(\tilde{\boldsymbol{x}}) > f(\boldsymbol{x}) \quad \text{and} \quad f^\circ(\tilde{\boldsymbol{x}}; \bar{\boldsymbol{x}} - \boldsymbol{x}) > 0.$$

Denote $\tilde{\boldsymbol{x}} = \tilde{\lambda}\boldsymbol{x} + (1 - \tilde{\lambda})\boldsymbol{y}$, where $\tilde{\lambda} \in (\bar{\lambda}, 1)$. From the definitions of points $\bar{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}$ we see that

$$\bar{\boldsymbol{x}} - \boldsymbol{x} = (1 - \bar{\lambda})(\boldsymbol{y} - \boldsymbol{x}) \quad \text{and} \quad \boldsymbol{y} - \tilde{\boldsymbol{x}} = \tilde{\lambda}(\boldsymbol{y} - \boldsymbol{x})$$

Thus,

$$\bar{\boldsymbol{x}} - \boldsymbol{x} = \frac{1 - \bar{\lambda}}{\tilde{\lambda}}(\boldsymbol{y} - \tilde{\boldsymbol{x}})$$

and

$$0 < f^\circ(\tilde{\boldsymbol{x}}; \bar{\boldsymbol{x}} - \boldsymbol{x}) = \frac{1-\bar{\lambda}}{\tilde{\lambda}} f^\circ(\tilde{\boldsymbol{x}}; \boldsymbol{y} - \tilde{\boldsymbol{x}}).$$

Thus, $0 < f^\circ(\tilde{\boldsymbol{x}}; \boldsymbol{y} - \tilde{\boldsymbol{x}})$ and $f(\tilde{\boldsymbol{x}}) > f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ which contradicts the l-quasiconvexity of function f. Hence, f is quasiconvex. □

COROLLARY 4.9. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz continuous and quasiconvex if and only if it is l-quasiconvex.*

PROOF. The result follows directly from Theorems 4.7 and 4.8. □

COROLLARY 4.10. *If f is f°-quasiconvex, then f is quasiconvex.*

PROOF. The result follows from Remark 4.2 and Theorem 4.8. □

Likewise the pseudomonotonicity there exists also a concept of quasimonotonicity (see [9]).

DEFINITION 4.11. The generalized directional derivative f° is called *quasimonotone*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0 \qquad \text{implies} \qquad f^\circ(\boldsymbol{y}; \boldsymbol{x} - \boldsymbol{y}) \leq 0$$

or, equivalently

$$\min \{f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}), f^\circ(\boldsymbol{y}; \boldsymbol{x} - \boldsymbol{y})\} \leq 0.$$

Note that analogously to the pseudomonotonicity (see Definition 3.6) we could define also the strict quasimonotonicity, but it would be equivalent to the pseudomonotonicity.

It turns out that the generalized directional derivative f° of the function $f : \mathbb{R}^n \to \mathbb{R}$ is quasimonotone if and only if the function is locally Lipschitz continuous and quasiconvex.

THEOREM 4.12. *If f° is quasimonotone, then f is quasiconvex.*

PROOF. Let us, on the contrary assume, that f is not quasiconvex. Then there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\bar{\lambda} \in (0, 1)$ such that

$$f(\bar{\boldsymbol{x}}) > f(\boldsymbol{x}) \geq f(\boldsymbol{y}),$$

where $\bar{\boldsymbol{x}} = \boldsymbol{x} + \bar{\lambda}(\boldsymbol{y} - \boldsymbol{x})$. Then by the Mean-Value Theorem 2.7 there exist $\hat{\boldsymbol{x}}, \tilde{\boldsymbol{x}} \in \mathbb{R}^n$ such that

$$f(\bar{\boldsymbol{x}}) - f(\boldsymbol{y}) \in \partial f(\hat{\boldsymbol{x}})^T(\bar{\boldsymbol{x}} - \boldsymbol{y})$$

and

$$f(\bar{\boldsymbol{x}}) - f(\boldsymbol{x}) \in \partial f(\tilde{\boldsymbol{x}})^T(\bar{\boldsymbol{x}} - \boldsymbol{x}),$$

where

$$\hat{\boldsymbol{x}} = \boldsymbol{x} + \hat{\lambda}(\boldsymbol{y} - \boldsymbol{x}), \quad \tilde{\boldsymbol{x}} = \boldsymbol{x} + \tilde{\lambda}(\boldsymbol{y} - \boldsymbol{x}), \quad 0 < \tilde{\lambda} < \bar{\lambda} < \hat{\lambda} < 1.$$

This means that, due to the definition of the Clarke subdifferential, there exist $\hat{\boldsymbol{\xi}} \in \partial f(\hat{\boldsymbol{x}})$ and $\tilde{\boldsymbol{\xi}} \in \partial f(\tilde{\boldsymbol{x}})$ such that

$$0 < f(\bar{\boldsymbol{x}}) - f(\boldsymbol{y}) = \hat{\boldsymbol{\xi}}^T(\bar{\boldsymbol{x}} - \boldsymbol{y}) \leq f^\circ(\hat{\boldsymbol{x}}; \bar{\boldsymbol{x}} - \boldsymbol{y}) = (1 - \bar{\lambda}) f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x} - \boldsymbol{y})$$

and

$$0 < f(\bar{\boldsymbol{x}}) - f(\boldsymbol{x}) = \tilde{\boldsymbol{\xi}}^T(\bar{\boldsymbol{x}} - \boldsymbol{x}) \leq f^\circ(\tilde{\boldsymbol{x}}; \bar{\boldsymbol{x}} - \boldsymbol{x}) = \bar{\lambda} f^\circ(\tilde{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x})$$

by the positive homogeneity of $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ (see Theorem 2.2 (i)). Then we deduce that

$$f^\circ(\hat{\boldsymbol{x}}; \tilde{\boldsymbol{x}} - \hat{\boldsymbol{x}}) = (\hat{\lambda} - \tilde{\lambda}) f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x} - \boldsymbol{y}) > 0$$

and

$$f^\circ(\tilde{\boldsymbol{x}}; \hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}}) = (\hat{\lambda} - \tilde{\lambda}) f^\circ(\tilde{\boldsymbol{x}}; \boldsymbol{y} - \boldsymbol{x}) > 0,$$

which contradicts the quasimonotonicity. Thus, f is quasiconvex. □

THEOREM 4.13. *If function $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz continuous and quasiconvex then the generalized directional derivative f° is quasimonotone.*

PROOF. On the contrary, assume that f° is not quasimonotone. Then there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0$ and $f^\circ(\boldsymbol{y}; \boldsymbol{x} - \boldsymbol{y}) > 0$ Let

$$\delta = \min \{f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}),\ f^\circ(\boldsymbol{y}; \boldsymbol{x} - \boldsymbol{y})\}.$$

Let $\varepsilon_1 > 0$ be such that the local Lipschitz condition holds in the ball $B(\boldsymbol{x}; \varepsilon_1)$ with Lipschitz constant K_1. Correspondingly, let $\varepsilon_2 > 0$ be such that the local Lipschitz condition holds in the ball $B(\boldsymbol{y}; \varepsilon_2)$ with Lipschitz constant K_2. Let $K = \max\{K_1, K_2\}$ and $\varepsilon = \min\left\{\frac{\delta}{4K}, \varepsilon_1, \varepsilon_2\right\}$. According to Theorem 2.10 there exists a sequence $(\boldsymbol{z}_1^i)$, such that f is differentiable, $\lim_{i\to\infty} \boldsymbol{z}_1^i = \boldsymbol{x}$ and an index $I \in \mathbb{N}$ such that

$$f'(\boldsymbol{z}_1^i; \boldsymbol{y} - \boldsymbol{x}) = \nabla f(\boldsymbol{z}_1^i)^T(\boldsymbol{y} - \boldsymbol{x}) \geq \frac{\delta}{2}$$

when $i \geq I$. Similarly, there exists a sequence $(\boldsymbol{z}_2^j)$, such that f is differentiable, $\lim_{j\to\infty} \boldsymbol{z}_2^j = \boldsymbol{y}$ and an index $J \in \mathbb{N}$ such that

$$f'(\boldsymbol{z}_2^j; \boldsymbol{x} - \boldsymbol{y}) = \nabla f(\boldsymbol{z}_2^j)^T(\boldsymbol{x} - \boldsymbol{y}) \geq \frac{\delta}{2}$$

when $j \geq J$. Let $\boldsymbol{z}_1 \in B(\boldsymbol{x}; \varepsilon) \cap \{(\boldsymbol{z}_1^i) \mid i \geq I\}$ and $\boldsymbol{z}_2 \in B(\boldsymbol{y}; \varepsilon) \cap \left\{(\boldsymbol{z}_2^j) \mid j \geq J\right\}$. Due to symmetry we may assume that $f(\boldsymbol{z}_1) \geq f(\boldsymbol{z}_2)$ without a loss of generality. According to Lemma 2.11

$$\begin{aligned} |f'(\boldsymbol{z}_1; \boldsymbol{z}_2 - \boldsymbol{z}_1) - f'(\boldsymbol{z}_1; \boldsymbol{y} - \boldsymbol{x})| &\leq K \|\boldsymbol{z}_2 - \boldsymbol{z}_1 - (\boldsymbol{y} - \boldsymbol{x})\| \\ \leq\ K\|\boldsymbol{x} - \boldsymbol{z}_1\| + K\|\boldsymbol{z}_2 - \boldsymbol{y}\| &< 2K\frac{\delta}{4K} = \frac{\delta}{2}. \end{aligned}$$

Since $f'(\boldsymbol{z}_1; \boldsymbol{y} - \boldsymbol{x}) > \frac{\delta}{2}$ also $f'(\boldsymbol{z}_1; \boldsymbol{z}_2 - \boldsymbol{z}_1) > 0$. Thus, there exists $\lambda \in (0, 1)$ such that

$$f(\boldsymbol{z}_1 + \lambda(\boldsymbol{z}_2 - \boldsymbol{z}_1)) > f(\boldsymbol{z}_1) \geq f(\boldsymbol{z}_2),$$

which contradicts the quasiconvexity. □

COROLLARY 4.14. *A function f is l-quasiconvex if and only if f° is quasimonotone.*

PROOF. The result follows from Corollary 4.9 and Theorems 4.12 and 4.13. □

COROLLARY 4.15. *If f is f°-quasiconvex, then f° is quasimonotone.*

PROOF. The results follows from Remark 4.2 and Corollary 4.14. □

By Corollary 4.10 f°-quasiconvex function is quasiconvex. The next result shows, that for a subdifferentially regular function quasiconvexity and f°-quasiconvexity coincides.

THEOREM 4.16. *If f is both quasiconvex and subdifferentially regular, then f is f°-quasiconvex.*

PROOF. Due to the subdifferential regularity f is locally Lipschitz continuous. Suppose, that $f(\boldsymbol{y}) \leq f(\boldsymbol{x})$. Then the subdifferential regularity and quasiconvexity implies, that

$$\begin{aligned} f^\circ(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) = f'(\boldsymbol{x};\boldsymbol{y}-\boldsymbol{x}) &= \lim_{t\downarrow 0} \frac{f(\boldsymbol{x}+t(\boldsymbol{y}-\boldsymbol{x}))-f(\boldsymbol{x})}{t} \\ &= \lim_{t\downarrow 0} \frac{f(t\boldsymbol{y}+(1-t)\boldsymbol{x})-f(\boldsymbol{x})}{t} \leq \lim_{t\downarrow 0} \frac{f(\boldsymbol{x})-f(\boldsymbol{x})}{t} = 0 \end{aligned}$$

in other words, f is f°-quasiconvex. □

COROLLARY 4.17. *A subdifferentially regular l-quasiconvex function is f°-quasiconvex.*

PROOF. The result follows from Corollary 4.9 and Theorem 4.16. □

COROLLARY 4.18. *A subdifferentially regular function f with quasimonotone f° is f°-quasiconvex.*

PROOF. The result follows from Corollaries 4.14 and 4.17. □

In Theorem 4.16 the subdifferential regularity cannot be omitted, as the next example shows.

EXAMPLE 4.1. Define $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) := \begin{cases} |x|, & x \in (-\infty, 1) \\ 1, & x \in [1,2] \\ x-1, & x \in (2,\infty). \end{cases}$$

Then f is clearly locally Lipschitz continuous and quasiconvex. However, by taking $x := 1$ and $y := 2$ we have $f^\circ(x; y-x) = f^\circ(1;1) = 1 > 0$, but $f(y) = f(2) = 1 \not> 1 = f(1) = f(x)$ and thus, due to Lemma 4.5, f is not f°-quasiconvex. Note that f is not subdifferentially regular since $f'(1;1) = 0 \neq 1 = f^\circ(1;1)$. Furthermore, f is not f°-pseudoconvex, since $0 \in \partial f(1) = [0,1]$ although $x = 1$ is not a global minimum (cf. Theorem 3.5).

As stated in Corollary 4.17 the subdifferential regularity ensures that the l-quasiconvexity implies f°-quasiconvexity. In sequel we show that also the following nonconstancy property [9] has the similar consequence.

DEFINITION 4.19. The function $f : \mathbb{R}^n \to \mathbb{R}$ is said to satisfy *nonconstancy property* (in short, NC-property), if there exists no line segment $[a, b]$ along which f is constant.

It is good to note that the subdifferential regularity and the NC-property are two separate concepts. An example of function which is subdifferentially regular but does not satisfy the NC-property is

$$g_1(x) = \begin{cases} (x+1)^2 & \text{, if } x \le -1 \\ 0 & \text{, if } -1 \le x \le 1 \\ (x-1)^2 & \text{, if } x \ge 1 \end{cases}.$$

On the other hand, the function

$$g_2(x) = \begin{cases} 2x & \text{, if } x \le 0 \\ \frac{1}{2}x & \text{, if } x \ge 0 \end{cases}$$

poses the NC-property but it is not subdifferentially regular since $g_2^\circ(0; 1) = 2 \neq \frac{1}{2} = g_2'(0; 1)$.

For the function with NC-property also the quasimonotonicity and the f°-quasiconvexity coincides.

THEOREM 4.20. *If f° is quasimonotone and f poses the NC-property, then f is f°-quasiconvex.*

PROOF. Let us, on the contrary, assume that f is not f°-quasiconvex. Then there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $f(\boldsymbol{y}) \le f(\boldsymbol{x})$ and

$$f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0. \tag{14}$$

According to Theorem 4.12 f is quasiconvex, which means that $f(\boldsymbol{z}) \le f(\boldsymbol{x})$ for all $\boldsymbol{z} \in [\boldsymbol{y}, \boldsymbol{x}]$. Thus, due to the NC-property, there exists $\bar\lambda \in (0, 1]$ such that $\bar{\boldsymbol{x}} = \boldsymbol{x} + \bar\lambda(\boldsymbol{y} - \boldsymbol{x})$ and $f(\bar{\boldsymbol{x}}) < f(\boldsymbol{x})$. By the Mean-Value Theorem 2.7 there exists $\hat{\boldsymbol{x}} \in (\bar{\boldsymbol{x}}, \boldsymbol{x})$ such that

$$f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}) \in \partial f(\hat{\boldsymbol{x}})^T(\boldsymbol{x} - \bar{\boldsymbol{x}}),$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x} + \hat\lambda(\bar{\boldsymbol{x}} - \boldsymbol{x})$ and $\hat\lambda \in (0, 1)$. This means that there exists $\hat{\boldsymbol{\xi}} \in \partial f(\hat{\boldsymbol{x}})$ such that

$$0 < f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}) = \hat{\boldsymbol{\xi}}^T(\boldsymbol{x} - \bar{\boldsymbol{x}}) \le f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x} - \bar{\boldsymbol{x}}). \tag{15}$$

On the other hand, from the positive homogeneity of $\boldsymbol{d} \mapsto f^\circ(\boldsymbol{x}; \boldsymbol{d})$ and (14) we deduce that

$$f^\circ(\boldsymbol{x}; \hat{\boldsymbol{x}} - \boldsymbol{x}) = \hat\lambda f^\circ(\boldsymbol{x}; \bar{\boldsymbol{x}} - \boldsymbol{x}) = \hat\lambda\bar\lambda f^\circ(\boldsymbol{x}; \boldsymbol{y} - \boldsymbol{x}) > 0.$$

Then the quasimonotonicity, the positive homogeneity and (15) imply that

$$0 \geq f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x} - \hat{\boldsymbol{x}}) = \hat{\lambda} f^\circ(\hat{\boldsymbol{x}}; \boldsymbol{x} - \bar{\boldsymbol{x}}) > 0,$$

which is impossible. Thus f is f°-quasiconvex. □

COROLLARY 4.21. *A l-quasiconvex function with NC-property is f°-quasiconvex.*

PROOF. The result follows from Corollary 4.14 and Theorem 4.20. □

COROLLARY 4.22. *A locally Lipschitz continuous and quasiconvex function with NC-property is f°-quasiconvex.*

PROOF. The result follows from Theorems 4.13 and 4.20. □

EXAMPLE 4.2. Consider the function in Example 4.1. Its generalized directional derivative is quasimonotone since the function is quasiconvex and locally Lipschitz continuous. However, the function does not satisfy the non-constancy property and, thus, it is not guaranteed to be f°-quasiconvex. As shown in Example 4.1 the function is not f°-quasiconvex.

The next results concerning the verification of the f°-quasiconvexity are, in practice, analogous to those of f°-pseudoconvexity.

LEMMA 4.23. *A locally Lipschitz continuous function $g : \mathbb{R} \to \mathbb{R}$ is increasing, if and only if $\varsigma \geq 0$ for all $\varsigma \in \partial g(x)$ and $x \in \mathbb{R}$.*

PROOF. The proof is almost similar to that of Lemma 3.11 by changing the meanings of $<$ and $\leq$. □

THEOREM 4.24. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be f°-quasiconvex and $g : \mathbb{R} \to \mathbb{R}$ locally Lipschitz continuous and increasing. Then the composite function $f := g \circ h : \mathbb{R}^n \to \mathbb{R}$ is also f°-quasiconvex.*

PROOF. The proof is similar to that of Theorem 3.12. □

THEOREM 4.25. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be f°-quasiconvex for all $i = 1, \ldots, m$. Then the function*

$$f(\boldsymbol{x}) := \max \{f_i(\boldsymbol{x}) \mid i = 1, \ldots, m\}$$

is also f°-quasiconvex.

PROOF. The proof is similar to that of Theorem 3.13. □

As in the case of f°-pseudoconvexity, the following property guarantees that the sum of f°-quasiconvex functions is also f°-quasiconvex.

DEFINITION 4.26. The functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are said to be *additively monotone*, if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda_i \geq 0$, $i = 1, \ldots, m$

$$\sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{y}) \leq \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) \qquad \text{implies} \qquad f_i(\boldsymbol{y}) \leq f_i(\boldsymbol{x}).$$

THEOREM 4.27. *Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be f°-quasiconvex and additively monotone, and $\lambda_i \geq 0$ for all $i = 1, \dots, m$. Then the function*

$$f(\boldsymbol{x}) := \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x})$$

is f°-quasiconvex.

PROOF. The proof is similar to that of Theorem 3.15. □

Finally we study the relations between pseudo- and quasiconvexity. According to [1] for differentiable functions pseudoconvexity implies quasiconvexity. Also, It turns out that f°-pseudoconvexity implies f°-quasiconvexity.

THEOREM 4.28. *An f°-pseudoconvex function is f°-quasiconvex.*

PROOF. On the contrary, assume that an f°-pseudoconvex function f is not f°-quasiconvex. Then, there exist points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $f^\circ(\boldsymbol{x}, \boldsymbol{y} - \boldsymbol{x}) > 0$ and $f(\boldsymbol{x}) = f(\boldsymbol{y})$. According to Lemma 3.9 this is impossible for f°-pseudoconvex function. Thus, f is f°-quasiconvex. □

COROLLARY 4.29. *If f° is pseudomonotone then it is also quasimonotone.*

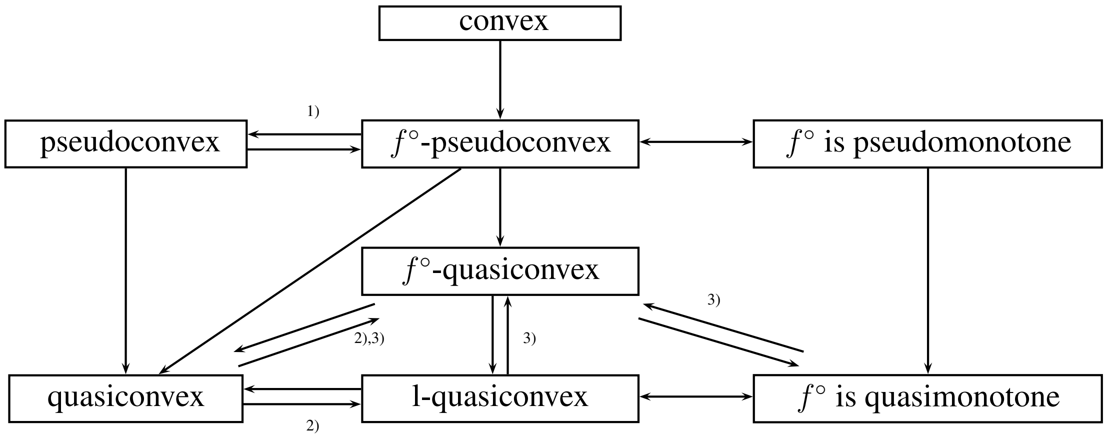
PROOF. The result follows from Corollary 4.15 and Theorems 3.7 and 4.28. □

The next example shows that the result in Theorem 4.28 cannot be converted.

EXAMPLE 4.3. Define $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) := x^3$. Clearly f is quasiconvex and as a smooth function also subdifferentially regular. Thus, by Theorem 4.16 it is f°-quasiconvex. However, by taking $x := 0$ and $y := -1$ we have $f^\circ(x; y - x) = f^\circ(0; -1) = 0$, but $f(y) = f(-1) = -1 \ngeq 0 = f(0) = f(x)$ and thus, due to Lemma 3.4, f is not f°-pseudoconvex.

5 Concluding Remarks

To the end we summarize all the relationships presented above:



1) demands continuous differentiability,

2) demands local Lipschitz continuity,

3) demands NC-property or subdifferential regularity.

Acknowledgements

We want to thank the University of Turku and the Magnus Ehrnrooth Foundation.

References

[1] AVRIEL, M., DIEWERT, W. E., SCHAIBLE, S., AND ZANG, I. *Generalized Concavity*. Plenum Press, New York, 1988.

[2] BHATIA, D., AND JAIN, P. Generalized (f,ρ)-convexity and duality for non smooth multi-objective programs. *Optimization 31* (1994), 153–164.

[3] BORDE, J., AND CROUZEIX, J. P. Continuity properties of the normal cone to the level sets of a quasiconvex function. *Journal of Optimization Theory and Applications 66* (1990), 415–429.

[4] CLARKE, F. H. *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York, 1983.

[5] DIEWERT, W. E. *Alternative Characterizations of Six Kinds of Quasiconcavity in the Nondifferentiable Case with Applications to Nonsmooth Programming*. ”Generalized Concavity in Optimization and Economics” (Eds. Schaible, S. and Ziemba, W. T.), Academic Press, New York, pp. 51–95.

[6] ELLAIA, R., AND HASSOUNI, A. Characterization of nonsmooth functions through their generalized gradients. *Optimization 22* (1991), 401–416.

[7] HIRIART-URRUTY, J. B. New concepts in nondifferentiable programming. *Bulletin de la Sociète Mathèmatiques de France, Mémoires 60* (1979), 57–85.

[8] KOMLÓSI, S. Some properties of nondifferentiable pseudoconvex functions. *Mathematical Programming 26* (1983), 232–237.

[9] KOMLÓSI, S. Generalized monotonicity and generalized convexity. *Journal of Optimization Theory and Applications 84* (1995), 361–376.

[10] MÄKELÄ, M. M., AND NEITTAANMÄKI, P. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co., Singapore, 1992.

[11] MANGASARIAN, O. L. Pseudoconvex functions. *SIAM Journal on Control 3* (1965), 281–290.

[12] NOBAKHTIAN, S. Infine functions and nonsmooth multiobjective optimization problems. *Computers and Mathematics with Applications 51* (2006), 1385–1394.

[13] NOBAKHTIAN, S. Multiobjective problems with nonsmooth equality constraints. *Numerical Functional Analysis and Optimization 30* (2009), 337–351.

[14] PINI, R., AND SINGH, C. A survey of recent [1985–1995] advances in generalized convexity with applications to duality theory and optimality conditions. *Optimization 39* (1997), 311–360.

[15] ROCKAFELLAR, R. T. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[16] SCHAIBLE, S. *Generalized Monotone Maps*. ”Nonsmooth Optimization: Methods and Applications” (Ed. Giannessi, F.), Gordon and Breach Science Publishers, Amsterdam, pp. 392–408.

[17] STAIB, T. Necessary optimality conditions for nonsmooth multicriteria optimization problem. *SIAM Journal on Optimization 2* (1992), 153–171.

[18] YANG, X. M., AND LIU, S. Y. Three kinds of generalized convexity. *Journal of Optimization Theory and Applications 86* (1995), 501–513.

A An example of differentiable but nonsmooth function

Here is the proof that function

$$f(x) = \begin{cases} 0, & x = 0 \\ x^2 \cos(\frac{1}{x}), & x \neq 0 \end{cases}$$

is locally Lipschitz continuous, differentiable but nonsmooth and there exists a point $y \in \mathbb{R}$ such that $\partial f(y) \neq \{\nabla f(y)\}$.

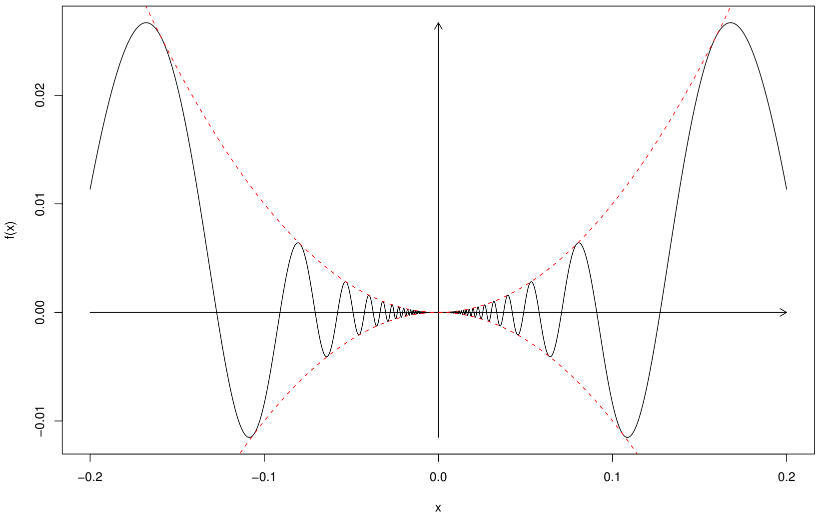


Figure 1: A plot of function f when $x \in [-0.2, 0.2]$.

We first show the differentiability. Function $\tilde{f}(x) := x^2 \cos(\frac{1}{x})$ is differentiable everywhere but at 0 and it's derivative is

$$\tilde{f}'(x) = \sin(\frac{1}{x}) + 2x\cos(\frac{1}{x}). \tag{16}$$

The derivative is also continuous when $x \neq 0$. Thus, f is continuously differentiable when $x \neq 0$. Since

$$f(0 + x) - f(0) = 0 \cdot x + |x - 0|\,|x| \cos(\frac{1}{x})$$

and $\lim_{x\to 0} |x| \cos(\frac{1}{x}) = 0$ the function f is differentiable at the point 0 and $f'(0) = 0$.

However, from (16) we see that the limit $\lim_{x\to 0} f'(x)$ does not exist implying that f is not continuously differentiable.

Next we prove that f is locally Lipschitz continuous. As stated before, continuously differentiable function is locally Lipschitz continuous. Hence, f is locally Lipschitz continuous, when $x \neq 0$. We prove that f is locally Lipschitz continuous also at the point $x = 0$ by considering Lipschitz condition with different values of $y, z \in (-1, 1), y \neq z$.

Let $-1 < y < z < 0$. The function f is continuously differentiable at (y, z). Then,

$$
\begin{aligned}
|f(z) - f(y)| &= \left| \int_y^z f'(x)dx \right| \leq \int_y^z \max_{x \in [y,z]} \{|f'(x)|\} \, dx \\
&= \max_{x \in [y,z]} \left\{ \left| \sin(\frac{1}{x}) + 2x \cos(\frac{1}{x}) \right| \right\} (z - y) \\
&\leq (1 + 2 \cdot 1 \cdot 1) |z - y| = 3 |z - y|,
\end{aligned}
$$

hence the Lipschitz condition holds. Due to the symmetry of the function f the Lipschitz condition holds also when $0 < y < z < 1$.

Now, let $-1 < y < 0$ and $0 < z < 1$. Then $|y + z| < |y - z|$ and the symmetry implies $f(-z) = f(z)$. Thus,

$$
|f(y) - f(z)| = |f(y) - f(-z)| \leq 3 |y + z| \leq 3 |y - z|,
$$

where the first inequality follows from the consideration of the case $-1 < y < z < 0$. Thus, the Lipschitz condition holds when $-1 < y < 0$ and $0 < z < 1$. Finally, let $y = 0$ and $z \in (-1, 1) \setminus \{0\}$. Then

$$
|f(0) - f(z)| = \left| z^2 \cos(\frac{1}{z}) \right| \leq |z| \, 1 \cdot 1 = |0 - z|,
$$

and the Lipschitz condition holds for this case too. Thus, the function f is locally Lipschitz continuous.

Consider the subdifferential of the function f at the point 0. By choosing the sequence $x^i = (\frac{1}{2i\pi + \frac{\pi}{2}})$, $i \in \mathbb{N}$ we see that $\lim_{i \to \infty} f'(x^i) = 1$. Correspondingly, by choosing the sequence $x^i = (\frac{1}{2i\pi - \frac{\pi}{2}})$, $i \in \mathbb{N}$ we see that $\lim_{i \to \infty} f'(x) = -1$. Thus, by Theorem 2.3

$$
[-1, 1] \subseteq \partial f(0).
$$

Particularly, $\partial f(0) \neq f'(0)$.

B An example of f°-pseudoconvex function with no directional derivative

Next we show a function that is f°-pseudoconvex, but whose directional derivative is not defined at every point.

Consider the following piecewise linear function

$$
f(x) = \begin{cases} x & \text{, if } x \leq 0 \\ 2^{(-1)^\alpha} \frac{1}{10^\alpha} + (\frac{5}{4} + (-1)^\alpha \frac{11}{12})(x - \frac{1}{10^\alpha}) & \text{, if } 0 < x < \frac{1}{10} \\ x - \frac{1}{20} & \text{, if } x \geq \frac{1}{10}, \end{cases}
$$

where

$$\alpha = \alpha(x) = \lfloor -\log_{10}(x) \rfloor .$$

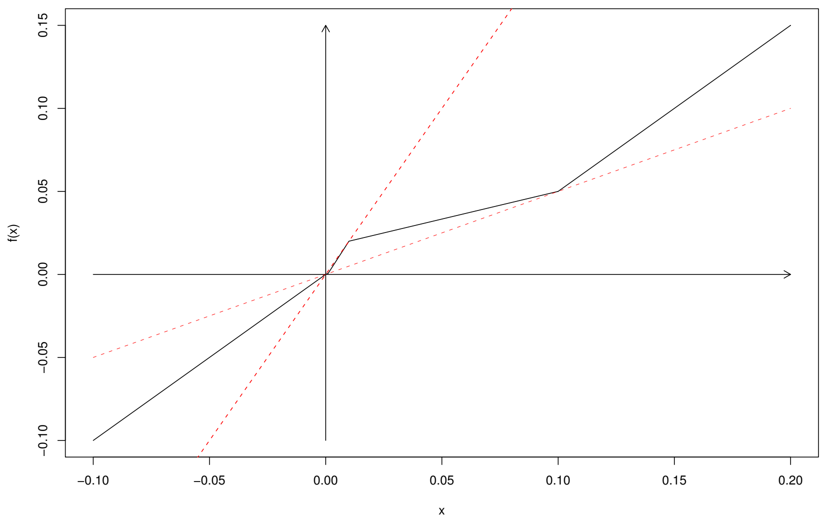


Figure 2: A plot of function f when $x \in [-0.1, 0.2]$.

The f is drawn in figure 2. The dashed lines represents lines $y = 2x$ and $y = \frac{1}{2}x$. Function f always lies between these two lines.

The function f is not differentiable at points $(\frac{1}{10^i})$, $i = 1, 2, \ldots$ and 0. From the definition of the function f we see that everywhere but at 0 the classical directional derivative $f'(x; 1)$ has an upper bound

$$\max\left\{1, \frac{5}{4} + \frac{11}{12}\right\} = \frac{5}{4} + \frac{11}{12} = \frac{13}{6}$$

and a lower bound

$$\min\left\{1, \frac{5}{4} - \frac{11}{12}\right\} = \frac{1}{3}.$$

Correspondingly, the directional derivative $f'(x; -1)$ has lower and upper bounds $-\frac{13}{6}$ and $-\frac{1}{3}$. When $x = 0$ wee see from the figure 2 that $|f(y)| \leq 2\,|y|$ for all $y \in \mathbb{R}$. Thus, we see that when $x, y \in \mathbb{R}$ the inequality

$$|f(x) - f(y)| \leq \frac{13}{6}\,|x - y|$$

holds and f is Lipschitz continuous. Actually, at an arbitrary point x^0 the function lies between the lines

$$y = \frac{13}{6}(x - x^0) + f(x^0) \quad \text{and} \quad y = \frac{1}{3}(x - x^0) + f(x^0).$$

Next, we prove the f°-pseudoconvexity of the function f. As stated previously, for all $x \in \mathbb{R}$, $t > 0$ the inequalities

$$-\frac{13}{6} \leq \frac{f(x-t) - f(x)}{t} \leq -\frac{1}{3},$$

holds implying $f^\circ(x, -1) \leq -\frac{1}{3}$. Now, the f°-pseudoconvexity follows from the fact that $f(y) < f(x)$ if and only if $y < x$.

Finally, we prove that directional derivative $f'(0; 1)$ does not exist.

THEOREM B.1. *Function f does not have the directional derivative $f'(0; 1)$.*

PROOF. Consider the limit

$$\lim_{t \downarrow 0} \varphi(t) = \lim_{t \downarrow 0} \frac{f(0+t) - f(0)}{t} \tag{17}$$

with different sequences (t^i). Let the sequence be $t^i = \frac{1}{10^{2i}}$, $i \in \mathbb{N}$. Then

$$\begin{aligned}
\alpha(t^i) &= 2i \\
f(t^i) &= 2\frac{1}{10^{2i}} + \frac{13}{6}(\frac{1}{10^{2i}} - \frac{1}{10^{2i}}) = 2\frac{1}{10^{2i}} \\
\varphi(t^i) &= \frac{2\frac{1}{10^{2i}}}{\frac{1}{10^{2i}}} = 2,
\end{aligned}$$

and the limit (17) is 2. Now, let the sequence be $s^i = \frac{1}{10^{2i+1}}$, $i \in \mathbb{N}$. Then

$$\begin{aligned}
\alpha(s^i) &= 2i + 1 \\
f(s^i) &= \frac{1}{2}\frac{1}{10^{2i+1}} + \frac{1}{3}(\frac{1}{10^{2i+1}} - \frac{1}{10^{2i+1}}) = \frac{1}{2}\frac{1}{10^{2i+1}} \\
\varphi(s^i) &= \frac{1}{2}\frac{\frac{1}{10^{2i+1}}}{\frac{1}{10^{2i+1}}} = \frac{1}{2},
\end{aligned}$$

and the limit (17) is $\frac{1}{2}$. The sequences (t^i) and (s^i) generates different limits and thus, the function f does not have the directional derivative $f'(0; 1)$. □

TURKU
CENTRE *for*
COMPUTER
SCIENCE

Lemminkäisenkatu 14 A, 20520 Turku, Finland | www.tucs.fi

University of Turku
- Department of Information Technology
- Department of Mathematics

Åbo Akademi University
- Department of Computer Science
- Institute for Advanced Management Systems Research

Turku School of Economics and Business Administration
- Institute of Information Systems Sciences

ISBN 978-952-12-2496-6
ISSN 1239-1891