Set-Valued Anal (2011) 19:75–96
DOI 10.1007/s11228-010-0136-x

Characterizing Convexity of a Function by Its Fréchet and Limiting Second-Order Subdifferentials

N. H. Chieu · T. D. Chuong · J.-C. Yao · N. D. Yen

Received: 27 October 2009 / Accepted: 2 March 2010 /
Published online: 20 March 2010

Abstract The Fréchet and limiting second-order subdifferentials of a proper lower semicontinuous convex function $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ have a property called the positive semi-definiteness (PSD)—in analogy with the notion of positive semi-definiteness of symmetric real matrices. In general, the PSD is insufficient for ensuring the convexity of an arbitrary lower semicontinuous function φ. However, if φ is a $C^{1,1}$ function then the PSD property of one of the second-order subdifferentials is a complete characterization of the convexity of φ. The same assertion is valid for C^1 functions of one variable. The limiting second-order subdifferential can recognize the convexity/nonconvexity of piecewise linear functions and of separable piecewise C^2 functions, while its Fréchet counterpart cannot.

This work was supported by National Sun Yat-Sen University (Kaohsiung, Taiwan) and National Foundation for Science & Technology Development (Vietnam). The authors would like to thank the referee for insightful comments and valuable suggestions.

N. H. Chieu
Department of Mathematics, Vinh University, Vinh, Nghe An, Vietnam
e-mail: nghuychieu@vinhuni.edu.vn, nghuychieu@gmail.com

T. D. Chuong
Department of Mathematics, Dong Thap University, Cao Lanh City, Dong Thap Province, Vietnam
e-mail: chuongthaidoan@yahoo.com

J.-C. Yao (✉)
Department of Applied Mathematics, National Sun Yat-Sen University, Kaohsiung 804, Taiwan
e-mail: yaojc@math.nsysu.edu.tw

N. D. Yen
Institute of Mathematics, Vietnamese Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi 10307, Vietnam
e-mail: ndyen@math.ac.vn

Keywords Convexity · Characterization · Positive semi-definite property · Fréchet second-order subdifferential · Limiting second-order subdifferential

Mathematics Subject Classifications (2000) 49K40 · 49J40 · 49J52 · 49J53

1 Introduction

Convexity of functions and sets plays a remarkable role in economics, management science, and mathematical theories (functional analysis, optimization theory, etc.). Therefore the study of convex functions and other concepts related to convexity are important from both the theoretical and practical points of view.

First-order characterizations for the convexity of extended-real-valued functions via the monotonicity of the Fréchet derivative (when it exists!) and the monotonicity of the Fréchet subdifferential mapping or the limiting subdifferential mapping can be found, e.g., in [7, 24] and [16, Theorem 3.56]. The convexity can be characterized also by using first-order directional derivatives; see e.g. [7] and the references therein.

The simplest and the most useful second-order characterization of convexity of real-valued functions is the theorem (see for instance [20, 24]) saying that a C^2 function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for every $x \in \mathbb{R}^n$ the Hessian $\nabla^2 f(x)$ is a positive semidefinite matrix. Relaxing the assumption on the C^2 smoothness of the function under consideration, several authors have characterized the convexity by using various kinds of generalized second-order directional derivatives. We refer the reader to [1, 4, 6, 9, 10, 25, 26] for many interesting results obtained in this direction.

Although the Fréchet and/or the limiting second-order subdifferential mappings have a significant role in variational analysis and its applications [5, 13–17, 19], it seems to us that their use in characterizing convexity of extended-real-valued functions has not been studied, so far. The purpose of this paper is to find out *to which extent the convexity can be characterized by these second-order subdifferential mappings*.

We will see that the Fréchet and limiting second-order subdifferentials of a proper lower semicontinuous convex function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}} = [-\infty, +\infty]$ have a property called the positive semi-definiteness (PSD)—in analogy with the notion of positive semi-definiteness of symmetric real matrices. In general, the PSD is insufficient for ensuring the convexity of an arbitrary lower semicontinuous function φ. However, if φ is a $C^{1,1}$ function then the PSD property of one of the second-order subdifferentials is a complete characterization of the convexity of φ. The same assertion is valid for C^1 functions of one variable. The limiting second-order subdifferential can recognize the convexity/nonconvexity of piecewise linear functions and of separable piecewise C^2 functions, while its Fréchet counterpart cannot. The obtained results are analyzed and illustrated by suitable examples.

Since strong convexity of functions and the related concept of strongly monotone operators have various applications in theory of algorithms (see e.g. [24]) and stability theory of optimization problems and variational inequalities (see e.g. [2, 11, 12, 27, 28]), by using the Fréchet and limiting second-order subdifferentials we will give some necessary and sufficient conditions for strong convexity.

The rest of the paper has five sections. Section 2 presents some basic definitions of variational analysis [16]. Section 3 describes a necessary condition for the convexity of a real-valued function by its Fréchet and limiting second-order subdifferentials. In Section 4, we establish second-order sufficient conditions for the convexity of real-valued functions from the following classes: (i) $C^{1,1}$ functions, (ii) C^1 functions of one variable, (iii) Piecewise linear functions, (iv) Separable piecewise C^2 functions. Characterizations of strong convexity are obtained in Section 5, while questions requiring further investigations are stated in Section 6.

2 Basic Definitions

We begin with some notions from [16] which will be needed in the sequel. For a set-valued mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$,

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{ x^* \in \mathbb{R}^n \Big|\ \exists \text{ sequences } x_k \to \bar{x} \text{ and } x_k^* \to x^* \text{ with } x_k^* \in F(x_k) \text{ for all } k \in \mathbb{N} \Big\}$$

denotes the *sequential Painlevé–Kuratowski upper limit* of F as $x \to \bar{x}$. The symbols $x \xrightarrow{\Omega} \bar{x}$ and $x \xrightarrow{\varphi} \bar{x}$ mean that $x \to \bar{x}$ with $x \in \Omega$ and $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$ for a set $\Omega \subset \mathbb{R}^n$ and an extended-real-valued function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$, respectively. Denote by $\mathcal{N}(x)$ the collection of the neighborhoods of $x \in \mathbb{R}^n$, by $\operatorname{cl} A$ the closure of $A \subset \mathbb{R}^n$, and by $\mathbb{B}(x, \varepsilon)$ the open ball centered at x with radius $\varepsilon > 0$.

Definition 2.1 Let $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in \mathbb{R}$ and let $\varepsilon \geqslant 0$. The $\varepsilon-$*subdifferential* of φ at $\bar{x}$ is the set $\widehat{\partial}_\varepsilon \varphi(\bar{x})$ defined by

$$\widehat{\partial}_\varepsilon \varphi(\bar{x}) := \left\{ x^* \in \mathbb{R}^n \mid \liminf_{x\to\bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geqslant -\varepsilon \right\}.$$

We put $\widehat{\partial}_\varepsilon \varphi(\bar{x}) = \emptyset$ if $|\varphi(\bar{x})| = \infty$. When $\varepsilon = 0$ the set $\widehat{\partial}_0 \varphi(\bar{x})$, denoted by $\widehat{\partial}\varphi(\bar{x})$, is called the *Fréchet subdifferential* of φ at $\bar{x}$. The *limiting subdifferential* (or *Mordukhovich subdifferential*) of φ at $\bar{x}$ is given by

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x\to\bar{x};\ \varepsilon\downarrow 0} \widehat{\partial}_\varepsilon \varphi(x),$$

where "Limsup" stands for the sequential Painlevé–Kuratowski upper limit of the set-valued mapping $F : \mathbb{R}^n \times [0, \infty) \rightrightarrows \mathbb{R}^n$ given by $F(x, \varepsilon) := \widehat{\partial}_\varepsilon \varphi(x)$.

It is clear that $\widehat{\partial}\varphi(\bar{x}) \subset \partial\varphi(\bar{x})$ and $\widehat{\partial}\varphi(\bar{x})$ is a closed convex set (may be empty). The examples given in [16] show that $\partial\varphi(\bar{x})$ is nonconvex in general (for instance, $\partial\varphi(0) = \{-1, 1\}$ for $\varphi(x) = -|x|,\ x \in \mathbb{R}$). One say that φ is *lower regular* at $\bar{x}$ if $\widehat{\partial}\varphi(\bar{x}) = \partial\varphi(\bar{x})$. If φ is convex and $\varphi(\bar{x})$ is finite, then [16] it holds

$$\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \left\{ x^* \in \mathbb{R}^n \mid \langle x^*, x - \bar{x}\rangle \leq \varphi(x) - \varphi(\bar{x})\ \forall x \in \mathbb{R}^n \right\}.$$

If φ is continuously differentiable at $\bar{x}$, then $\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$ (see [16]). Thus, convex functions and C^1 functions are lower regular at any point in their effective domains.

Definition 2.2 The *Fréchet normal cone* and the *Mordukhovich normal cone* to $\Omega \subset \mathbb{R}^n$ at x are defined, respectively, by $\widehat{N}(x; \Omega) := \widehat{\partial}\delta(x; \Omega)$ and $N(x; \Omega) := \partial\delta(x; \Omega)$, where $\delta(x; \Omega) = 0$ if $x \in \Omega$ and $\delta(x; \Omega) = \infty$ if $x \notin \Omega$ is the indicator function of Ω.

From the definition it follows that $x^* \in \widehat{N}(x; \Omega)$ if and only if

$$\limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \leq 0.$$

Let $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping with the *graph*

$$\text{gph}\, F := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in F(x)\}.$$

Definition 2.3 The *Mordukhovich coderivative* $D^*F(\bar{x}, \bar{y}) \colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and the *Fréchet coderivative* $\widehat{D}^*F(\bar{x}, \bar{y}) \colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ of F at $(\bar{x}, \bar{y}) \in \text{gph}\, F$ are defined respectively by

$$D^*F(\bar{x}, \bar{y})(y^*) := \left\{x^* \in \mathbb{R}^n \middle| \, (x^*, -y^*) \in N((\bar{x}, \bar{y}); \text{gph}\, F)\right\}$$

and

$$\widehat{D}^*F(\bar{x}, \bar{y})(y^*) := \left\{x^* \in \mathbb{R}^n \middle| \, (x^*, -y^*) \in \widehat{N}((\bar{x}, \bar{y}); \text{gph}\, F)\right\}.$$

We omit $\bar{y} = f(\bar{x})$ in the above coderivative notions if $F = f \colon \mathbb{R}^n \to \mathbb{R}^m$ is single-valued.

If $f \colon \mathbb{R}^n \to \mathbb{R}^m$ is *strictly differentiable* at $\bar{x}$ in the sense that

$$\lim_{x,u \to \bar{x}} \frac{f(x) - f(u) - \langle \nabla f(\bar{x}), x - u\rangle}{\|x - u\|} = 0,$$

where $\nabla f(\bar{x})$ is the Fréchet derivative of f at $\bar{x}$, then

$$D^*f(\bar{x})(y^*) = \widehat{D}^*f(\bar{x})(y^*) = \left\{\nabla f(\bar{x})^* y^*\right\} \text{ for all } \; y^* \in \mathbb{R}^m.$$

These formulae show that the coderivatives under consideration are two appropriate extensions of the adjoint derivative operator of real-valued functions to the case of set-valued maps.

Having a rich calculus supported by effective characterizations of Lipschitzian and related properties of set-valued mappings, the coderivatives have become one of the most remarkable notions in modern variational analysis; see [16, 21] and references therein. One can use coderivatives to construct the second-order generalized differential theory of extended-real-valued functions. Such approaches were initiated in Mordukhovich [14].

Definition 2.4 Let $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function with a finite value at $\bar{x}$.

(i) For any $\bar{y} \in \partial\varphi(\bar{x})$, the map $\partial^2\varphi(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with the values

$$\partial^2\varphi(\bar{x}, \bar{y})(u) = (D^*\partial\varphi)(\bar{x}, \bar{y})(u) \quad (u \in \mathbb{R}^n)$$

is said to be the *limiting second-order subdifferential* of φ at $\bar{x}$ relative to $\bar{y}$.

(ii) For any $\bar{y} \in \widehat{\partial}\varphi(\bar{x})$, the map $\widehat{\partial}^2\varphi(\bar{x}, \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with the values

$$\widehat{\partial}^2\varphi(\bar{x}, \bar{y})(u) = (\widehat{D}^*\widehat{\partial}\varphi)(\bar{x}, \bar{y})(u) \quad (u \in \mathbb{R}^n)$$

is said to be the *Fréchet second-order subdifferential* of φ at $\bar{x}$ relative to $\bar{y}$.

Clearly, if φ is lower regular at any point in a neighborhood of $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \widehat{\partial}\varphi(\bar{x})$, then $\widehat{\partial}^2\varphi(\bar{x}, \bar{y})(u) \subset \partial^2\varphi(\bar{x}, \bar{y})(u)$ for all $u \in \mathbb{R}^n$.

If φ is a C^2 function around $\bar{x}$ (i.e., φ is twice continuously differentiable in a neighborhood of the point) then

$$\partial^2\varphi(\bar{x})(u) = \widehat{\partial}^2\varphi(\bar{x})(u) = \left\{(\nabla^2\varphi(\bar{x}))^* u\right\} \quad \forall u \in \mathbb{R}^n,$$

where $(\nabla^2\varphi(\bar{x}))^*$ is the adjoint operator of the Hessian $\nabla^2\varphi(\bar{x})$. Various properties and calculus rules for the limiting second-order subdifferential can be found in [16, 17].

Definition 2.5 (see [18] and [21, Chap. 12]) One says that a set-valued map $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a *monotone operator* if

$$\langle x^* - y^*, x - y\rangle \geq 0 \quad \text{for all } x, y \in \mathbb{R}^n,\ x^* \in T(x),\ y^* \in T(y).$$

A monotone operator T is said to be *maximal monotone* if its graph gph$T$ is not a proper subset of the graph of any other monotone operator.

We refer to [18, 20, 21] for detailed information on maximal operators and their applications.

In analogy with positive semi-definiteness and positive definiteness of real matrices, one can consider the following concepts.

Definition 2.6 A set-valued map $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *positive semi-definite* (PSD for brevity) if $\langle z, u\rangle \geq 0$ for any $u \in \mathbb{R}^n$ and $z \in T(u)$. If $\langle z, u\rangle > 0$ whenever $u \in \mathbb{R}^n \setminus \{0\}$ and $z \in T(u)$, then T is said to be *positive definite*.

In [19, Theorem 1.3], Poliquin and Rockafellar have proved that the positive definiteness of the limiting second-order subdifferential mapping $\partial^2\varphi(\bar{x}, 0) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ characterizes the *tilt stability* of a stationary point $\bar{x}$ of a function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ (provided that φ has some required properties). Later, Levy, Poliquin and Rockafellar [13, Theorem 2.3] have shown that the positive definiteness of a parametric limiting second-order subdifferential mapping can be used to study the *full stability* [13, Def. 1.1] of local optimal points. For the sake of completeness, we observe that by [21, Corollary 8.47(a)] the set $\partial\varphi(\bar{x})$ of the (general) subgradients of φ at $\bar{x}$ defined in [19, p. 288] coincides with the limiting subdifferential $\partial\varphi(\bar{x})$ in Definition 2.1. Note in addition that the *subgradient mapping* of φ in [13, p. 583] is the same as the limiting subdifferential mapping $x \mapsto \partial\varphi(x)$ given by Definition 2.1.

3 Second-Order Necessary Conditions

Theorem 3.2 below gives a *necessary condition* for the convexity of a function $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ via its limiting second-order subdifferential mapping. For proving it, we shall need an auxiliary result.

Lemma 3.1 ([19, Theorem 2.1]) *If $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone operator, then for every point $(x, y) \in$ gph$T$ it holds*

$$\langle z, u\rangle \geq 0 \quad \textit{whenever } z \in D^*T(x, y)(u),$$

i.e., the coderivative mapping $D^*T(x, y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is positive semi-definite.*

Theorem 3.2 *Let* $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be proper lower semicontinuous. If* φ *is convex, then*

$$\langle z, u\rangle \geq 0 \ \ \textit{for all } u \in \mathbb{R}^n \textit{ and } z \in \partial^2\varphi(x, y)(u) \textit{ with } (x, y) \in \operatorname{gph}\partial\varphi; \tag{3.1}$$

that is, for every $(x, y) \in \operatorname{gph}\partial\varphi$*, the mapping* $\partial^2\varphi(x, y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is positive semi-definite.*

Proof If φ is convex then, according to [21, Theorem 12.17], $\partial\varphi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone operator. Applying Lemma 3.1 to $T := \partial\varphi$ yields

$$\langle z, u\rangle \geq 0 \quad \text{for all } u \in \mathbb{R}^n,\ z \in D^*\partial\varphi(x, y)(u) \text{ with } (x, y) \in \operatorname{gph}\partial\varphi.$$

Since $\partial^2\varphi(x, y)(u) = D^*\partial\varphi(x, y)(u)$ by definition, Eq. 3.1 holds. □

Remark 3.3 Suppose that $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ is a proper lower semicontinuous convex function. Then $\widehat{\partial}^2\varphi(x, y)(u) \subset \partial^2\varphi(x, y)(u)$ and $\partial^2\varphi(x, y)(\cdot)$ is PSD; hence $\widehat{\partial}^2\varphi(x, y)(\cdot)$ is also PSD:

$$\langle z, u\rangle \geq 0 \ \text{ for all } u \in \mathbb{R}^n,\ z \in \widehat{\partial}^2\varphi(x, y)(u) \text{ with } (x, y) \in \operatorname{gph}\widehat{\partial}\varphi. \tag{3.2}$$

Thus, from Theorem 3.2 it follows that Eq. 3.2 is a second-order *necessary condition* for the convexity of φ.

In general, the PSD property of both the second-order subdifferentials $\partial^2\varphi(\cdot)$ and $\widehat{\partial}^2\varphi(\cdot)$ of a proper lower semicontinuous $\varphi : \mathbb{R}^n \to \mathbb{R}$ does not imply the convexity of φ.

Example 3.4 Let $\varphi(x) = 0$ for $x \in \mathbb{R} \setminus \{0\}$ and $\varphi(x) = -1$ for $x = 0$. Note that $\varphi : \mathbb{R} \to \mathbb{R}$ is lower semicontinuous on $\mathbb{R}$ and

$$\partial\varphi(x) = \widehat{\partial}\varphi(x) = \begin{cases} \{0\} & \text{if } x \neq 0 \\ \mathbb{R} & \text{if } x = 0. \end{cases}$$

Therefore, for each $(x, y) \in \operatorname{gph}\partial\varphi$, we have

$$\partial^2\varphi(x, y)(u) = \begin{cases} \{z \in \mathbb{R} \mid (z, -u) \in \{0\} \times \mathbb{R}\} & \text{if } x \neq 0,\ y = 0 \\ \{z \in \mathbb{R} \mid (z, -u) \in (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})\} & \text{if } x = 0,\ y = 0 \\ \{z \in \mathbb{R} \mid (z, -u) \in \mathbb{R} \times \{0\}\} & \text{if } x = 0,\ y \neq 0. \end{cases}$$

It follows that $zu \geq 0$ for any $u \in \mathbb{R}$ and $z \in \partial^2\varphi(x, y)(u)$ with $(x, y) \in \operatorname{gph}\partial\varphi$. Thus Eq. 3.1 is valid. Of course, Eq. 3.2 is also fulfilled. Nevertheless, φ is nonconvex.

We have seen that neither Eq. 3.1 nor Eq. 3.2 can guarantee the convexity of φ. In other words, the class of the proper lower semicontinuous functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ is too large for one's checking their convexity by analyzing the second-order subdifferentials $\widehat{\partial}^2\varphi(\cdot)$ and $\partial^2\varphi(\cdot)$.

It is interesting, however, that the PSD property of the second-order subdifferential $\widehat{\partial}^2\varphi(\cdot)$ (or $\partial^2\varphi(\cdot)$) can serve as a *sufficient condition* for the convexity if one considers smaller classes of functions. This will be clear from the results of the forthcoming section.

4 Second-Order Sufficient Conditions

4.1 $C^{1,1}$ Functions

Let $\mathcal{L}^m$ be the $\sigma-$algebra of the Lebesgue measurable subsets on $\mathbb{R}^m$ $(m = 1, 2, ...)$. The $\sigma-$algebra product (see [3, 22]) of $\mathcal{L}^1$ and $\mathcal{L}^{n-1}$ is denoted by $\mathcal{L}^1 \times \mathcal{L}^{n-1}$. Let λ and ν, respectively, stand for the Lebesgue measures on $\mathbb{R}$ and on $\mathbb{R}^{n-1}$. We denote the product measure (see [3, 22]) of λ and ν by $\lambda \times \nu$, and the Lebesgue measure on $\mathbb{R}^n$ by μ.

Lemma 4.1 *Suppose that $\Omega_0 \in \mathcal{L}^n$, $\mu(\mathbb{R}^n \backslash \Omega_0) = 0$, and $a, b \in \mathbb{R}^n$ are distinct points. Then there exist sequences of vectors $a^k \to a$ and $b^k \to b$ such that*

$$\lambda\left(\left\{t \in [0,1] \mid a^k + t(b^k - a^k) \in \Omega_0\right\}\right) = 1 \quad (\forall k \in \mathbb{N}).$$

Proof Put $\tilde{e}^1 := \|b-a\|^{-1}(b-a)$ and choose $\{\tilde{e}^i\}_{i=2}^n$ such that $\{\tilde{e}^1, \tilde{e}^2, ..., \tilde{e}^n\}$ is an orthonormal basis of $\mathbb{R}^n$. Denote by Λ the linear operator from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying $\Lambda(e^i) = \tilde{e}^i$, $i = 1, 2, ..., n$, where e^i is the $i-$th unit vector in $\mathbb{R}^n$. If $a = \sum_{i=1}^{n} \alpha_i \tilde{e}^i$ then

$$b = a + \|b-a\|\tilde{e}^1 = (\alpha_1 + \|b-a\|)\tilde{e}^1 + \sum_{i=2}^{n} \alpha_i \tilde{e}^i.$$

For each $k \in \mathbb{N}$, let

$$A_k := \left\{x = \sum_{i=1}^{n} x_i \tilde{e}^i \mid x_1 \in [\alpha_1, \alpha_1 + \|b-a\|],\ x_i \in \left[\alpha_i, \alpha_i + k^{-1}\right],\ i = 2, 3, ..., n\right\}.$$

We have

$$\Lambda^{-1}(A_k) = [\alpha_1, \alpha_1 + \|b-a\|] \times \left[\alpha_2, \alpha_2 + k^{-1}\right] \times ... \times \left[\alpha_n, \alpha_n + k^{-1}\right].$$

Since Λ is a linear isometry, $\mu(\Lambda^{-1}(A)) = \mu(A)$ for any $A \in \mathcal{L}^n$; see [23, Theorem 5.2, p. 140]. In particular,

$$\mu\left(\mathbb{R}^n \backslash \Lambda^{-1}(\Omega_0)\right) = \mu\left(\Lambda^{-1}(\mathbb{R}^n \backslash \Omega_0)\right) = \mu(\mathbb{R}^n \backslash \Omega_0) = 0.$$

Hence

$$\mu(\tilde{M}) = \mu(U \times V) = k^{1-n}\|b-a\|,$$

where $U = [\alpha_1, \alpha_1 + \|b-a\|]$, $V = \left[\alpha_2, \alpha_2 + k^{-1}\right] \times ... \times \left[\alpha_n, \alpha_n + k^{-1}\right]$, and $\tilde{M} = (U \times V) \cap \Lambda^{-1}(\Omega_0)$. Since μ is the completion of the product measure $\lambda \times \nu$ (see [22, Theorem 7.11]), there exists $M \in \mathcal{L}^1 \times \mathcal{L}^{n-1}$ satisfying $M \subset \tilde{M}$ and $\mu(M) = \mu(\tilde{M})$. Since $\mu(M) = (\lambda \times \nu)(M)$, we have

$$(\lambda \times \nu)(M) = \mu(\tilde{M}) = k^{1-n}\|b-a\|. \tag{4.1}$$

Define $M_v = \{u \in \mathbb{R} \mid (u, v) \in M\}$, $f(u, v) = \chi_M(u, v)$ where $\chi_M(\cdot)$ denotes the characteristic function of M, and $f_v(u) = f(u, v)$ for all $(u, v) \in U \times V$. Note that $f_v(u) = \chi_{M_v}(u)$. We can find $v^k \in V$ such that $\lambda(M_{v^k}) = \|b - a\|$. Indeed, by the Fubini theorem ([22, Theorem 7.8],[3, Theorem 5.2.2])

$$(\lambda \times \nu)(M) = \int_{\mathbb{R}^n} f(u, v) d(\lambda \times \nu)(u, v) = \int_V d\nu(v) \int_U f_v(u) d\lambda(u)$$

$$= \int_V d\nu(v) \int_{M_v} d\lambda(u) = \int_V \lambda(M_v) d\nu(v)$$

$$\leq \|b - a\| \nu(V) = k^{1-n} \|b - a\|.$$

Thus, by Eq. 4.1, the last inequality must hold as an equality. Since $\lambda(M_v) \leq \|b - a\|$ for all $v \in V$, this implies $\lambda(M_v) = \|b - a\|$ for almost all $v \in V$. Let

$$a^k := \Lambda(\alpha_1, v^k) \quad \text{and} \quad b^k := \Lambda(\alpha_1 + \|b - a\|, v^k).$$

Then

$$\lambda\left(\left\{t \in [0, 1] \mid a^k + t(b^k - a^k) \in \Omega_0\right\}\right) = \lambda\left(\left\{t \in [0, 1] \mid \Lambda(\alpha_1 + t\|b - a\|, v^k) \in \Omega_0\right\}\right)$$

$$\geq \|b - a\|^{-1} \lambda(M_{v^k}) = 1.$$

It follows that $\lambda\left(\left\{t \in [0, 1] \mid a^k + t(b^k - a^k) \in \Omega_0\right\}\right) = 1$. As

$$v^k \in V = \left[\alpha_2, \alpha_2 + k^{-1}\right] \times ... \times \left[\alpha_n, \alpha_n + k^{-1}\right],$$

it holds $\lim_{k \to \infty} v^k = (\alpha_2, ..., \alpha_n)$. Hence $a^k \to a$ and $b^k \to b$. □

As usual [16, p. 123], we say that $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a $C^{1,1}$ function if φ is Fréchet differentiable and its derivative $\nabla\varphi(\cdot)$ is locally Lipschitz. The class of $C^{1,1}$ functions encompasses the class of C^2 functions. The formula $\varphi(x) = x|x|$ defines a $C^{1,1}$ function on $\mathbb{R}$, which is not a C^2 function.

In combination with Remark 3.3, the next statement shows that the PSD property of the Fréchet second-order subdifferential mapping provides a complete characterization for convexity of $C^{1,1}$ functions.

Theorem 4.2 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,1}$ function. If the condition 3.2 is fulfilled, then φ is convex.*

Proof Suppose that φ is a $C^{1,1}$ function and $\langle z, u \rangle \geq 0$ for all $u \in \mathbb{R}^n$, $z \in \widehat{\partial}^2\varphi(x, y)(u)$ with $(x, y) \in \text{gph}\widehat{\partial}\varphi$. Since φ is Fréchet differentiable, $\widehat{\partial}\varphi(x) = \{\nabla\varphi(x)\}$ for all $x \in \mathbb{R}^n$.

If φ is twice differentiable at x, then $\nabla^2\varphi(x)^*u \in \widehat{\partial}^2\varphi(x, \nabla\varphi(x))(u)$ for every $u \in \mathbb{R}^n$. Indeed, we have

$$\begin{aligned}
&\limsup_{\tilde{x}\to x} \frac{\langle \nabla^2\varphi(x)^*u, \tilde{x}-x\rangle - \langle u, \nabla\varphi(\tilde{x}) - \nabla\varphi(x)\rangle}{\|\tilde{x}-x\| + \|\nabla\varphi(\tilde{x}) - \nabla\varphi(x)\|} \\
&\le \limsup_{\tilde{x}\to x} \max\left\{0, \frac{\langle \nabla^2\varphi(x)^*u, \tilde{x}-x\rangle - \langle u, \nabla\varphi(\tilde{x}) - \nabla\varphi(x)\rangle}{\|\tilde{x}-x\|}\right\} \\
&= \limsup_{\tilde{x}\to x} \max\left\{0, \frac{\langle -u, \nabla\varphi(\tilde{x}) - \nabla\varphi(x) - \nabla^2\varphi(x)(\tilde{x}-x)\rangle}{\|\tilde{x}-x\|}\right\} \\
&= 0.
\end{aligned}$$

This means that $\nabla^2\varphi(x)^*u \in \widehat{D}^*\nabla\varphi(x)(u)$. Hence

$$\nabla^2\varphi(x)^*u \in \widehat{\partial}^2\varphi(x, \nabla\varphi(x))(u) \tag{4.2}$$

for all $u \in \mathbb{R}^n$, provided that φ is twice differentiable at x. Put

$$\Omega_0 = \{x \in \mathbb{R}^n \mid \varphi \text{ is twice differentiable at } x\}.$$

Since $\nabla\varphi(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous, by the Rademacher theorem we have $\Omega_0 \in \mathcal{L}^n$ and $\mu(\mathbb{R}^n \setminus \Omega_0) = 0$. We will prove that

$$\langle \nabla\varphi(y) - \nabla\varphi(x), y - x\rangle \ge 0 \quad \forall x, y \in \mathbb{R}^n. \tag{4.3}$$

It can be supposed that $\nabla\varphi$ is Lipschitz with modulus ℓ on $\bar{I\!B}([x, y]; 2)$. By Lemma 4.1 there exist sequences $x^k \to x$ and $y^k \to y$ such that

$$\lambda\left(\left\{t \in [0, 1] \mid x^k + t(y^k - x^k) \in \Omega_0\right\}\right) = 1.$$

Without loss of generality we can assume that $x^k, y^k \in \bar{I\!B}([x, y]; 2)$ for all k. Applying the Newton–Leibniz formula to $f(t) := \langle \nabla\varphi(x^k + t(y^k - x^k)), y^k - x^k\rangle$, we get

$$\begin{aligned}
\left\langle \nabla\varphi(y^k) - \nabla\varphi(x^k), y^k - x^k\right\rangle &= \int_0^1 f'(t)dt \\
&= \int_{T^k} (y^k - x^k)^T \nabla^2\varphi(x^k + t(y^k - x^k))(y^k - x^k)dt,
\end{aligned}$$

where $T^k := \{t \in [0, 1] \mid x^k + t(y^k - x^k) \in \Omega_0\}$. For any $t \in T^k$,

$$\nabla^2\varphi(x^k + t(y^k - x^k))(y^k - x^k) \in \widehat{\partial}^2\varphi(x^k + t(y^k - x^k))(y^k - x^k)$$

by Eq. 4.2; hence $(y^k - x^k)^T \nabla^2\varphi(x^k + t(y^k - x^k))(y^k - x^k) \ge 0$ by our assumption. Consequently,

$$\left\langle \nabla\varphi(y^k) - \nabla\varphi(x^k), y^k - x^k\right\rangle \ge 0 \quad \forall k.$$

Letting $k \to \infty$ we get Eq. 4.3. This means that $\nabla\varphi(\cdot)$ is monotone. Hence φ is convex. □

Remark 4.3 An equivalent formulation of Theorem 4.2 had appeared in [8] where the Fréchet second-order subdifferential $\widehat{\partial}^2\varphi$ is replaced by the so-called generalized Hessian. Note that no explicit proof was given in [8]. The proof hints in [8, Example 2.2] are based on the second-order Taylor expansion while the above proof appeals the Newton–Leibniz formula. The two approaches are quite different.

4.2 C^1 Functions of One Variable

We now establish a refined version of Theorem 4.2 in the case $n = 1$ which shows that, for C^1 functions of one variable, the PSD property of the Fréchet second-order subdifferential mapping implies convexity of the given function. (Hence the PSD property of the limiting second-order subdifferential mapping also implies convexity of the given function.)

Theorem 4.4 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$, $n = 1$, be a C^1 function. If the condition 3.2 is fulfilled, then φ is convex.*

Proof On the contrary, suppose that one could find a C^1 function $\varphi : \mathbb{R} \to \mathbb{R}$ satis fying

$$zu \geq 0 \quad \text{for all } u \in \mathbb{R},\ z \in \widehat{\partial}^2\varphi(x, y)(u) \ \text{ with } (x, y) \in \operatorname{gph}\widehat{\partial}\varphi, \tag{4.4}$$

but φ is nonconvex. Setting $f(x) = \nabla\varphi(x)$ and noting that $\widehat{\partial}\varphi(x) = \{f(x)\}$, we can use the definitions of Fréchet second-order subdifferential and Fréchet coderivative to rewrite Eq. 4.4 equivalently as follows

$$zu \leq 0 \quad \text{for all } (z, u) \in \widehat{N}((x, f(x)); \operatorname{gph} f) \ \text{ with } x \in \mathbb{R}. \tag{4.5}$$

By [16, Theorem 3.56], since φ is nonconvex, the Fréchet subdifferential mapping $\widehat{\partial}\varphi(\cdot) = \{f(\cdot)\}$ is nonmonotone. This means that there exists a pair $a, b \in \mathbb{R}$ such that $(f(b) - f(a))(b - a) < 0$. In order to obtain a fact contradicting to Eq. 4.5, we will use the following geometrical constructions: 1. Project orthogonally the curve $\Gamma := \big\{(x, f(x)) : x \in [a, b]\big\} \subset \operatorname{gph} f$ to the straight line $L := \big\{\lambda\big(f(a) - f(b), b - a\big) : \lambda \in \mathbb{R}\big\}$ which is orthogonal to the segment $\big\{(1 - t)\big(a, b(a)\big) + t\big(b, g(b)\big) : t \in [0, 1]\big\}$ connecting the chosen "bad points" $\big(a, b(a)\big), \big(b, g(b)\big) \in \operatorname{gph} f$; 2. Find the maximal (or minimal) value λ of the obtained projections on L and a point $\bar{z} \in \Gamma$ corresponding to this value; 3. Define a Fréchet normal to $\operatorname{gph} f$ at $\bar{z}$.

There is no loss of generality in assuming that $b - a > 0$ and $f(b) - f(a) < 0$. Put $w = \big(f(a) - f(b), b - a\big)$,

$$\psi(x) := \langle w, (x, f(x))\rangle = \big(f(a) - f(b)\big)x + (b - a) f(x)$$

and consider the optimization problems

$$\psi(x) \to \max \quad \text{s.t. } \ x \in [a, b] \tag{4.6}$$

and

$$\psi(x) \to \min \quad \text{s.t. } \ x \in [a, b]. \tag{4.7}$$

Since $\psi(a) = \psi(b)$ and ψ is continuous on $[a, b]$, at least one of the problems Eq. 4.6, Eq. 4.7 must have a global solution on the open interval (a, b).

First, suppose that Eq. 4.6 possesses a global solution $\bar{x} \in (a, b)$. Setting $\bar{z} = \big(\bar{x},\ f(\bar{x})\big)$, we want to show that

$$w \in \widehat{N}(\bar{z}; \operatorname{gph} f). \tag{4.8}$$

By the choice of $\bar{x}$, for any $x \in (a, b)$ it holds

$$\begin{aligned} 0 \geq \psi(x) - \psi(\bar{x}) &= \big(f(a) - f(b)\big)(x - \bar{x}) + (b - a)\big(f(x) - f(\bar{x})\big) \\ &= \langle w, (x, f(x)) - \bar{z}\rangle. \end{aligned}$$

This implies

$$\limsup_{z \xrightarrow{\operatorname{gph} f} \bar{z}} \frac{\langle w, z - \bar{z}\rangle}{\|z - \bar{z}\|} \leq 0,$$

hence Eq. 4.8 is valid. Since both the components of w are positive, Eq. 4.8 is in conflict with Eq. 4.5.

Now suppose that Eq. 4.7 admits a global solution $\bar{x} \in (a, b)$. Setting $\bar{z} = \big(\bar{x},\ f(\bar{x})\big)$, by an argument similar to the above we can show that

$$-w \in \widehat{N}(\bar{z}; \operatorname{gph} f). \tag{4.9}$$

Since both the components of $-w$ are negative, Eq. 4.9 contradicts Eq. 4.5.

Thus, in each of the two possible cases, we have obtained a contradiction. The proof is complete. □

The geometrical constructions used for proving Theorem 4.4 can hardly be applicable to the case $n \geq 2$. The question about whether or not the conclusion of that theorem holds for $n \neq 1$ remains open (see the last section).

4.3 Limiting Second-Order Subdifferential Mapping vs Fréchet Second-Order Subdifferential Mapping

The PSD property of the Fréchet second-order subdifferential mapping is not subtle enough to recognize the nonconvexity of a Lipschitz function.

Example 4.5 Let $\varepsilon > 0$ and let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined by

$$\varphi(x) = \begin{cases} 0 & \text{if } x \in (-\infty, -\varepsilon] \cup [\varepsilon, +\infty) \\ -\frac{1}{\varepsilon}x - 1 & \text{if } x \in (-\varepsilon, 0] \\ \frac{1}{\varepsilon}x - 1 & \text{if } x \in (0, \varepsilon). \end{cases}$$

It is easy to verify that

$$\partial\varphi(x) = \begin{cases} \{0\} & \text{if } x \in (-\infty, -\varepsilon) \cup (\varepsilon, +\infty) \\ \{0, -\frac{1}{\varepsilon}\} & \text{if } x = -\varepsilon \\ \{-\frac{1}{\varepsilon}\} & \text{if } x \in (-\varepsilon, 0) \\ [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] & \text{if } x = 0 \\ \{\frac{1}{\varepsilon}\} & \text{if } x \in (0, \varepsilon) \\ \{0, \frac{1}{\varepsilon}\} & \text{if } x = \varepsilon. \end{cases}$$

For $(x, y) = (\varepsilon, 0) \in \mathrm{gph}\partial\varphi$, we find that

$$\partial^2\varphi(x, y)(u) = \partial^2\varphi(\varepsilon, 0)(u) = \{z \in \mathbb{R} \mid (z, -u) \in (-\infty, 0] \times \mathbb{R}\} \quad \forall u \in \mathbb{R}.$$

Since $zu < 0$ for any $u > 0$ and $z \in \partial^2\varphi(\varepsilon, 0) \setminus \{0\}$, the subdifferential mapping $\partial^2\varphi(\varepsilon, 0)(\cdot)$ is not PSD. Meanwhile,

$$\widehat{\partial}\varphi(x) = \begin{cases} \{0\} & \text{if } x \in (-\infty, -\varepsilon) \cup (\varepsilon, +\infty) \\ \{\emptyset\} & \text{if } x = -\varepsilon \\ \{-\frac{1}{\varepsilon}\} & \text{if } x \in (-\varepsilon, 0) \\ [-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] & \text{if } x = 0 \\ \{\frac{1}{\varepsilon}\} & \text{if } x \in (0, \varepsilon) \\ \{\emptyset\} & \text{if } x = \varepsilon \end{cases}$$

and

$$\widehat{\partial}^2\varphi(x, y)(u) = \begin{cases} \{z \in \mathbb{R} \mid (z, -u) \in \{0\} \times \mathbb{R}\} & \text{if } |x| > \varepsilon,\ y = 0 \\ \{z \in \mathbb{R} \mid (z, -u) \in \{0\} \times \mathbb{R}\} & \text{if } 0 < x < \varepsilon,\ y = \frac{1}{\varepsilon} \\ \{z \in \mathbb{R} \mid (z, -u) \in \{0\} \times \mathbb{R}\} & \text{if } -\varepsilon < x < 0,\ y = -\frac{1}{\varepsilon} \\ \{z \in \mathbb{R} \mid (z, -u) \in \mathbb{R}_+ \times \mathbb{R}_-\} & \text{if } x = 0,\ y = -\frac{1}{\varepsilon} \\ \{z \in \mathbb{R} \mid (z, -u) \in \mathbb{R}_- \times \mathbb{R}_+\} & \text{if } x = 0,\ y = \frac{1}{\varepsilon} \\ \{z \in \mathbb{R} \mid (z, -u) \in \mathbb{R} \times \{0\}\} & \text{if } x = 0,\ -\frac{1}{\varepsilon} < y < \frac{1}{\varepsilon} \end{cases}$$

for every $(x, y) \in \mathrm{gph}\widehat{\partial}\varphi$ and $u \in \mathbb{R}$. Hence $zu \geq 0$ for all $u \in \mathbb{R}$, $z \in \widehat{\partial}^2\varphi(x, y)(u)$ with $(x, y) \in \mathrm{gph}\widehat{\partial}\varphi$. Although φ is Lipschitz and nonconvex, the Fréchet second-order subdifferential mapping of φ has the PSD property Eq. 3.2.

As the last example showed, the limiting second-order subdifferential mapping of φ does not have the PSD property (Eq. 3.1). In other words, although the nonconvexity of φ remains unnoticed by examining the Fréchet second-order subdifferential mapping, one can still recognize the nonconvexity of the function by checking the PSD property of its limiting second-order subdifferential mapping. Thus, the latter might be used for obtaining sufficient conditions for convexity of locally Lipschitz functions (or even of continuous functions), while its counterpart, the Fréchet second-order subdifferential mapping, cannot be used for the purpose.

4.4 Piecewise Linear Functions

A function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is said to be *piecewise linear* (or *piecewise affine*) if there exist families $\{P_1, ..., P_k\}$, $\{a_1, ..., a_k\}$, and $\{b_1, ..., b_k\}$ of polyhedral convex sets in $\mathbb{R}^n$, points in $\mathbb{R}^n$, and points in $\mathbb{R}$, respectively, such that $\mathbb{R}^n = \bigcup_{i=1}^{k} P_i$, $\mathrm{int}P_i \cap \mathrm{int}P_j = \emptyset$ for all $i \neq j$, and

$$\varphi(x) = \varphi_i(x) := \langle a_i, x\rangle + b_i \quad \forall x \in P_i,\ \forall i \in \{1, ..., k\}. \tag{4.10}$$

From Eq. 4.10 it follows that $\varphi_i(x) = \varphi_j(x)$ whenever $x \in P_i \cap P_j$ and $i, j \in \{1, ..., k\}$.

Lemma 4.6 *Let* A *and* B *be two closed subsets of* $\mathbb{R}^n$. *If* $\mathrm{int}A = \mathrm{int}B = \emptyset$, *then* $\mathrm{int}(A \cup B) = \emptyset$. *If* $\mathrm{int}A = \emptyset$, *then* $\mathrm{int}B = \mathrm{int}(A \cup B)$.

Proof To prove the first claim, suppose on the contrary that $\text{int}A = \text{int}B = \emptyset$, but $\text{int}(A \cup B) \neq \emptyset$. For any $x \in \text{int}(A \cup B)$, there are two cases: (i) $x \in B$, (ii) $x \in A$. Let the second case happen. Take any $U \in \mathcal{N}(x)$. Clearly, there is $V \in \mathcal{N}(x)$ satisfying $V \subset U$ and $V \subset \text{int}(A \cup B)$. Since $\text{int}A = \emptyset$ and $V \subset A \cup B$, we must have $V \cap B \neq \emptyset$. So, $U \cap B \neq \emptyset$ for all $U \in \mathcal{N}(x)$. This means that $x \in \text{cl}B = B$. Thus, in both cases, $x \in B$. We have shown that $\text{int}(A \cup B) \subset B$. This contradicts the fact that $\text{int}B = \emptyset$.

To prove the second claim, we first note that $\text{int}B \subset \text{int}(A \cup B)$. To obtain the reverse inclusion under the assumption $\text{int}A = \emptyset$, we take any $x \in \text{int}(A \cup B)$. Arguing as in the proof of the first claim, we have $U \cap B \neq \emptyset$ for all $U \in \mathcal{N}(x)$. Hence $x \in \text{cl}B = B$. This shows that $\text{int}(A \cup B) \subset B$, and thus $\text{int}(A \cup B) \subset \text{int}B$. □

Lemma 4.7 *For* $I := \{i \in \{1, 2, ..., k\} \mid \text{int}P_i \neq \emptyset\}$, *it holds* $\bigcup_{i \in I} P_i = \mathbb{R}^n$.

Proof Put $J = \{1, 2, ..., k\} \backslash I$. The first assertion of Lemma 4.6 implies that $\text{int}(\bigcup_{j \in J} P_j) = \emptyset$. Set $A = \bigcup_{j \in J} P_j$ and $B = \bigcup_{i \in I} P_i$. Since P_i $(i = 1, \dots, k)$ are closed, according to the second claim of Lemma 4.6, $\text{int}B = \text{int}(A \cup B)$. It follows that $\text{int}(\bigcup_{i \in I} P_i) = \text{int}(\bigcup_{i=1}^{k} P_i) = \mathbb{R}^n$. Hence $\bigcup_{i \in I} P_i = \mathbb{R}^n$. □

Lemma 4.8 *Let* $[x, y]$ *be an interval in* $\mathbb{R}^n$ $(x \neq y)$, $0 = \tau_0 < \tau_1 < ... < \tau_{m-1} < \tau_m = 1$ $(m \in \mathbb{N},\ m > 1)$, *and* $x_i := x + \tau_i(y - x)$ $(i = 0, 1, ..., m)$. *Suppose that* φ *is nonconvex and continuous on* $[x, y]$. *Then there must exist* $i \in \{0, 1, ..., m-2\}$ *such that* φ *is nonconvex on* $[x_i, x_{i+2}]$.

Proof Replacing φ by the function $g(t) := \varphi(x + t(y - x))$ $(t \in \mathbb{R})$ if necessary, we can assume that $n = 1$ and $[x, y] = [0, 1]$ (thus $x_i = \tau_i$, $i = 0, 1, ..., m$). Moreover, it suffices to prove the claim for the case $m = 3$, because the case $m = 2$ is trivial, and the case $m > 3$ can be treated by induction. To obtain a contradiction, suppose that φ is nonconvex on $[\tau_0, \tau_3]$ but it is convex on each one of the intervals $[\tau_0, \tau_2]$ and $[\tau_1, \tau_3]$. Take any $t_1, t_2 \in (0, 1) = (\tau_0, \tau_3)$, $t_1 < t_2$, and $\xi_i \in \partial\varphi(t_i)$ $(i = 1, 2)$. We will show that $\xi_1 \leq \xi_2$. If $t_2 \in (\tau_0, \tau_2)$ or $t_1 \in (\tau_1, \tau_3)$ then $\xi_1 \leq \xi_2$ by the convexity of φ on each one of the intervals $[\tau_0, \tau_2]$ and $[\tau_1, \tau_3]$ (the latter implies the monotonicity of the subdifferential mapping $\partial\varphi(\cdot)$ on the corresponding interval). Suppose that $t_1 \in (\tau_0, \tau_1]$ and $t_2 \in [\tau_2, \tau_3)$. Taking any $t \in (\tau_1, \tau_2)$ and $\xi \in \partial\varphi(t)$, we get $\xi_1 \leq \xi$ and $\xi \leq \xi_2$ by the above proof; so $\xi_1 \leq \xi_2$. Thus $\partial\varphi(\cdot)$ is monotone on (τ_0, τ_3), hence φ is convex on (τ_0, τ_3). This fact and the continuity of φ on $[\tau_0, \tau_3]$ imply that φ is convex on $[\tau_0, \tau_3]$, a contradiction. □

Theorem 4.9 *Let* $\varphi : \mathbb{R}^n \to \mathbb{R}$ *be a piecewise linear function. If Eq. 3.1 holds, then* φ *is convex.*

Proof According to Lemma 4.7, we can assume that $\text{int}P_i \neq \emptyset$ for all $i = 1, 2, ..., k$. Suppose that Eq. 3.1 holds, but φ is nonconvex. Consider the following two cases.

Case 1: $k = 2$. Suppose that $P_1 = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \le \alpha\}$, $P_2 = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \ge \alpha\}$, $a \neq 0$, $P_{12} = P_1 \cap P_2$, and

$$\varphi(x) = \begin{cases} \langle a_1, x\rangle + b_1 & \text{if } x \in P_1 \\ \langle a_2, x\rangle + b_2 & \text{if } x \in P_2, \end{cases}$$

where $a_1, a_2 \in \mathbb{R}^n$ $(a_1 \neq a_2)$ and $\langle a_1, x\rangle + b_1 = \langle a_2, x\rangle + b_2$ for all $x \in P_{12}$. Observe that $\mathbb{R}^n$ is the union of the disjoint sets $\operatorname{int} P_1$, $\operatorname{int} P_2$, and P_{12}. We now compute the limiting subdifferential $\partial\varphi(x)$.

If $x \in \operatorname{int} P_1$, then $\partial\varphi(x) = \widehat{\partial}\varphi(x) = \{a_1\}$.

If $x \in \operatorname{int} P_2$, then $\partial\varphi(x) = \widehat{\partial}\varphi(x) = \{a_2\}$.

Let $x \in P_{12}$. Since φ is convex on each one of the sets P_1 and P_2 but it is nonconvex on $\mathbb{R}^n = P_1 \cup P_2$, we can find $x_0 \in \operatorname{int} P_1$, $y_0 \in \operatorname{int} P_2$, and $t_1 \in (0, 1)$, such that

$$\varphi(z_1) > (1 - t_1)\varphi(x_0) + t_1\varphi(y_0), \tag{4.11}$$

where $z_1 := (1 - t_1)x_0 + t_1 y_0$. Let $t_0 \in (0, 1)$ be such that $z_0 := (1 - t_0)x_0 + t_0 y_0 \in P_{12}$. We will show that

$$\varphi(z_0) > (1 - t_0)\varphi(x_0) + t_0\varphi(y_0). \tag{4.12}$$

If $t_0 = t_1$ then Eq. 4.12 follows from Eq. 4.11, because $z_0 = z_1$. If $t_0 \in (0, t_1)$ then $z_1 = (1 - \lambda)y_0 + \lambda z_0$ with $\lambda = (1 - t_1)/(1 - t_0) \in (0, 1)$. Since φ is affine on $[z_0, y_0] \subset P_2$,

$$\varphi(z_1) = (1 - \lambda)\varphi(y_0) + \lambda\varphi(z_0).$$

Then, from Eq. 4.11 it follows that

$$\begin{aligned}\varphi(z_0) &> \lambda^{-1}[(1 - t_1)\varphi(x_0) + t_1\varphi(y_0) - (1 - \lambda)\varphi(y_0)] \\ &= (1 - t_0)\varphi(x_0) + t_0\varphi(y_0).\end{aligned}$$

This establishes Eq. 4.12. Similarly, Eq. 4.12 also holds for the case where $t_0 \in (t_1, 1)$. Since $x_0 \in P_1$, $y_0 \in P_2$, $z_0 \in P_{12}$,

$$\varphi(x_0) = \langle a_1, x_0\rangle + b_1, \quad \varphi(y_0) = \langle a_2, y_0\rangle + b_2$$

and

$$\varphi(z_0) = \langle a_1, z_0\rangle + b_1 = \langle a_2, z_0\rangle + b_2.$$

Hence, Eq. 4.12 implies

$$(1 - t_0)(\langle a_1, z_0\rangle + b_1) + t_0(\langle a_2, z_0\rangle + b_2) > (1 - t_0)(\langle a_1, x_0\rangle + b_1) + t_0(\langle a_2, y_0\rangle + b_2)$$

or, equivalently,

$$(1 - t_0)\langle a_1, z_0 - x_0\rangle + t_0\langle a_2, z_0 - y_0\rangle > 0.$$

As $z_0 = (1 - t_0)x_0 + t_0 y_0$, we obtain

$$\langle a_1 - a_2, y_0 - x_0\rangle > 0. \tag{4.13}$$

We are going to prove that $\widehat{\partial}\varphi(x) = \emptyset$. To the contrary suppose that there exists $x^* \in \widehat{\partial}\varphi(x)$. Then

$$\liminf_{u \to x} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x\rangle}{\|u - x\|} \ge 0. \tag{4.14}$$

It is clear that $u_j := x - \frac{1}{j}(y_0 - x_0) \to x$ as $j \to \infty$. Since $\langle a, x_0\rangle < \alpha$, $\langle a, y_0\rangle > \alpha$, and $\langle a, x\rangle = \alpha$, it holds

$$\langle a, u_j\rangle = \langle a, x\rangle - \tfrac{1}{j}\langle a, y_0 - x_0\rangle$$

$$= \alpha - \tfrac{1}{j}[\langle a, y_0\rangle - \langle a, x_0\rangle] < \alpha.$$

This means that $u_j \in P_1$ for all $j \in I\!N$. By Eq. 4.14,

$$\liminf_{j\to\infty} \frac{\left\langle a_1, x - \frac{1}{j}(y_0 - x_0)\right\rangle + b_1 - (\langle a_1, x\rangle + b_1) - \left\langle x^*, -\frac{1}{j}(y_0 - x_0)\right\rangle}{\frac{1}{j}\|y_0 - x_0\|} \geq 0.$$

Hence

$$\langle a_1, y_0 - x_0\rangle \leq \langle x^*, y_0 - x_0\rangle.$$

Similarly, choosing $u'_j = x + \frac{1}{j}(y_0 - x_0)$ we have $u'_j \to x$ and $u'_j \in P_2$ for all $j \in \mathbb{N}$. Substituting $u = u'_j$ into Eq. 4.14 yields

$$\langle x^*, y_0 - x_0\rangle \leq \langle a_2, y_0 - x_0\rangle.$$

Thus, we obtain $\langle a_1 - a_2, y_0 - x_0\rangle \leq 0$, a contradiction to Eq. 4.13. Therefore, $\widehat{\partial}\varphi(x) = \emptyset$ for all $x \in P_{12}$. We have

$$\begin{aligned}\partial\varphi(x) &= \operatorname*{Lim\,sup}_{u\to x} \widehat{\partial}\varphi(u)\\ &= \operatorname*{Lim\,sup}_{u\xrightarrow{\operatorname{int}P_1}x} \widehat{\partial}\varphi(u) \cup \operatorname*{Lim\,sup}_{u\xrightarrow{\operatorname{int}P_2}x} \widehat{\partial}\varphi(u) \cup \operatorname*{Lim\,sup}_{u\xrightarrow{P_{12}}x} \widehat{\partial}\varphi(u)\\ &= \{a_1, a_2\}.\end{aligned}$$

Hence

$$\partial\varphi(x) = \begin{cases}\{a_1\} & \text{if } x \in \operatorname{int}P_1,\\ \{a_2\} & \text{if } x \in \operatorname{int}P_2,\\ \{a_1, a_2\} & \text{if } x \in P_{12}.\end{cases}$$

By a simple computation we find

$$\partial^2\varphi(x, y)(u) = \begin{cases}\{0\} & \text{if } x \in \operatorname{int}P_1 \cup \operatorname{int}P_2 \text{ and } y \in \partial\varphi(x),\\ \mathbb{R}_+ a & \text{if } x \in P_{12} \text{ and } y = a_1,\\ \mathbb{R}_- a & \text{if } x \in P_{12} \text{ and } y = a_2\end{cases}$$

for all $u \in \mathbb{R}^n$. Taking $x \in P_{12}$, $y = a_1$, $z = a$ and $u = -a$, we get $z \in \partial^2\varphi(x, y)(u)$ and $\langle z, u\rangle < 0$. This contradicts Eq. 3.1. We have thus shown that φ is a convex function.

Case 2: $k > 2$. Since φ is assumed to be a nonconvex function, there exist $x, y \in \mathbb{R}^n$ such that $x \neq y$ and φ is nonconvex on the line segment $[x, y]$. As φ is piecewise linear, there exist $0 = \tau_0 < \tau_1 < ... < \tau_{m-1} < \tau_m = 1$ ($m \in I\!N, m > 1$), and $x_i := x + \tau_i(y - x)$ ($i = 0, 1, ..., m$) such that φ is affine on each interval $[x_i, x_{i+1}]$ ($i = 0, 1, ..., m - 1$). By Lemma 4.8 there exists $i \in \{0, 1, ..., m - 2\}$ such that φ is nonconvex on $[x_i, x_{i+2}]$. Hence, without loss of generality, we can assume that there exists $\bar{x} \in (x, y)$ such that φ is affine on each one of the segments $[x, \bar{x}]$ and $[\bar{x}, y]$ but it is nonconvex on $[x, y]$. Let $t \in (0, 1)$ be such that $\bar{x} = (1-t)x + ty$. From our assumption it follows that

$$\varphi(\bar{x}) > (1-t)\varphi(x) + t\varphi(y). \tag{4.15}$$

For every $u \in \mathbb{R}^n$, we set $I(u) = \{i \in \{1, 2, ..., k\} \mid u \in P_i\}$. Choose $\varepsilon > 0$ as small as $I\!B(\bar{x}, \varepsilon) \cap P_i = \emptyset$ for all $i \in \{1, 2, ..., k\} \backslash I(\bar{x})$. We can assume that $x, y \in I\!B(\bar{x}, \varepsilon)$. It is not difficult to see that $|I(\bar{x})| \geq 2$ and $\bar{x} \notin \mathrm{int} P_i$ for all i.

If $|I(\bar{x})| = 2$ then, by using the result of Case 1, we obtain a contradiction.

If $|I(\bar{x})| > 2$ then $\dim L \leq n-2$ with $L := \mathrm{aff}(\bigcap_{i \in I(\bar{x})} P_i)$ denoting the affine hull of $\bigcap_{i \in I(\bar{x})} P_i$. Indeed, without loss of generality we may assume that $\{1, 2, 3\} \subset I(\bar{x})$. Since $\mathrm{int} P_i \neq \emptyset$, $\mathrm{int} P_j \neq \emptyset$ and $\mathrm{int} P_i \cap \mathrm{int} P_j = \emptyset$, by the separation theorem we find a hyperplane L_{ij} separating the sets $\mathrm{intP_i}$ and $\mathrm{int} P_j$ $(1 \leq i < j \leq 3)$. Since the situation $L_{12} = L_{13} = L_{23}$ cannot occur, $\dim(L_{12} \cap L_{13} \cap L_{23}) \leq n-2$. Noting that $L \subset (L_{12} \cap L_{13} \cap L_{23})$, we have $\dim L \leq n-2$. In the case where $y \in L$, by invoking the last property we can find $\tilde{y} \in \mathbb{R}^n \setminus L$ as close to y as desired. Define $\tilde{x}_{\tilde{y}}$ by the condition $\bar{x} = (1-t)\tilde{x}_{\tilde{y}} + t\tilde{y}$. Clearly, $\tilde{x}_{\tilde{y}} \notin L$ and $\tilde{x}_{\tilde{y}} \to x$ as $\tilde{y} \to y$. Therefore, according to Eq. 4.15 we can find $\tilde{y} \in \mathbb{R}^n \setminus L$ such that

$$\varphi(\bar{x}) > (1-t)\varphi(\tilde{x}_{\tilde{y}}) + t\varphi(\tilde{y}).$$

Thus, replacing $\{x, y\}$ by $\{\tilde{x}_{\tilde{y}}, \tilde{y}\}$ if necessary, we can assume that $y \notin L$ and $x \notin L$. (Note that such replacement may destroy the property of φ of being affine on each one of the segments $[x, \bar{x}]$ and $[\bar{x}, y]$. But this property will not be employed in the sequel.)

Choose $\rho > 0$ as small as $I\!B(y, \rho) \subset I\!B(\bar{x}, \varepsilon)$, $I\!B(y, \rho) \cap L = \emptyset$, $x \notin I\!B(y, \rho)$, and φ is nonconvex on $[x, z]$ for each $z \in I\!B(y, \rho)$. We are going to show that there exists $z \in I\!B(y, \rho)$ such that $[x, z] \cap L = \emptyset$. Arguing by contradiction, we suppose that $[x, z] \cap L \neq \emptyset$ for all $z \in I\!B(y, \rho)$. Let us choose $y_1, y_2, ..., y_{n-1} \in I\!B(y, \rho)$ such that the system $\{x - y, y_1 - y, ..., y_{n-1} - y\}$ is linearly independent. For every $i \in \{1, \ldots, n-1\}$, there is some $\bar{x}_i \in (x, y_i) \cap L$. Clearly, for each i there exist real numbers α_i and β_i satisfying $\bar{x}_i - \bar{x} = \alpha_i(x - y) + \beta_i(y_i - y)$. It is not difficult to show that $\beta_i \neq 0$ for all i. Due to this fact and the linear independence of $\{x - y, y_1 - y, ..., y_{n-1} - y\}$, the system $\{\bar{x}_1 - \bar{x}, ..., \bar{x}_{n-1} - \bar{x}\} \subset L - \bar{x}$ is linearly independent. Hence we can assert that $\dim L \geq n-1$, a contradiction. We have thus proved that there exists $z \in I\!B(y, \rho)$ with $[x, z] \cap L = \emptyset$. Note that $[x, z] \subset I\!B(\bar{x}, \varepsilon)$ and φ is nonconvex on $[x, z]$. By Lemma 4.8 we can find $[x', y'] \subset [x, z]$ and $\bar{x}' \in (x', y')$ such that φ is convex on each one of the segments $[x', \bar{x}']$ and $[\bar{x}', y']$, and it is nonconvex on $[x', y']$. Since $\bar{x}' \in I\!B(\bar{x}, \varepsilon) \backslash [\bigcap_{i \in I(\bar{x})} P_i]$ and $I\!B(\bar{x}, \varepsilon) \cap P_i = \emptyset$ for all $i \in \{1, 2, ..., k\} \backslash I(\bar{x})$, we can assert that $|I(\bar{x}')| < |I(\bar{x})|$. Hence, if $|I(\bar{x})| > 2$ then we can find $[x', y']$ and $\bar{x}' \in (x', y')$ with $|I(\bar{x}')| < |I(\bar{x})|$ such that φ is convex on each one of the segments $[x', \bar{x}']$ and $[\bar{x}', y']$, but it is nonconvex on $[x', y']$. Repeating this procedure finitely many times we will come to the situation $|I(\bar{x})| = 2$ treated before. The proof is complete. □

4.5 Piecewise C^2 Functions of a Special Type

Following [17], we consider an extended-real-valued function $\varphi : \mathbb{R} \to \bar{\mathbb{R}}$ having the form

$$\varphi(\cdot) = \vartheta(\cdot) + \delta(\cdot, \Gamma), \tag{4.16}$$

where $\delta(\cdot, \Gamma)$ is the indicator function of a closed interval $\Gamma := [\alpha, \beta] \subset \bar{\mathbb{R}}$ (the possibilities $\alpha = -\infty$ and $\beta = +\infty$ are not excluded) and ϑ is *piecewise* C^2 in the following sense:

(i) ϑ is continuous on an open set $\mathcal{O}$ containing Γ;
(ii) there exist points $\kappa^1, \kappa^2, ..., \kappa^k$ in Γ with $\alpha < \kappa^1 < \kappa^2 < ... < \kappa^k < \beta$ and C^2 functions $\vartheta^j : \mathcal{O} \to \mathbb{R}$ ($j = 0, 1, ..., k$) such that

$$\vartheta(\xi) = \begin{cases} \vartheta^0(\xi) & \text{for } \xi \in [\alpha, \kappa^1], \\ \vartheta^j(\xi) & \text{for } \xi \in [\kappa^j, \kappa^{j+1}], \ \ j = 1, 2, ..., k-1, \\ \vartheta^k(\xi) & \text{for } \xi \in [\kappa^k, \beta]. \end{cases}$$

Define

$$M = \left\{ j \in \{1, 2, ...k\} \mid \nabla\vartheta^{j-1}(\kappa^j) \leq \nabla\vartheta^j(\kappa^j) \right\}$$

and

$$\mathcal{A}^j(\xi) = \left\{ (w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^j(\xi)z = 0 \right\}.$$

Take any $(\bar{p}, \bar{v}) \in \operatorname{gph}\partial\varphi$. According to [17],

$$N((\bar{p}, \bar{v}); \operatorname{gph}\partial\varphi) = \begin{cases} \mathcal{A}^0(\bar{p}) & \text{if } \bar{p} \in (\alpha, \kappa^1), \\ \mathcal{A}^j(\bar{p}) & \text{if } \bar{p} \in (\kappa^j, \kappa^{j+1}), \quad j = 1, 2, ..., k-1, \\ \mathcal{A}^k(\bar{p}) & \text{if } \bar{p} \in (\kappa^k, \beta). \end{cases}$$

For $\bar{p} = \alpha$,

$$N((\alpha, \bar{v}); \operatorname{gph}\partial\varphi) = \begin{cases} \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \quad \text{if } \bar{v} < \nabla\vartheta^0(\alpha), \\ \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \cup \mathcal{A}^0(\alpha) \\ \cup\{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^0(\alpha)z \leq 0, z \geq 0\} \\ \qquad\qquad\qquad \text{if } \bar{v} = \nabla\vartheta^0(\alpha). \end{cases}$$

For $\bar{p} = \beta$,

$$N((\beta, \bar{v}); \operatorname{gph}\partial\varphi) = \begin{cases} \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \quad \text{if } \bar{v} > \nabla\vartheta^k(\beta), \\ \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \cup \mathcal{A}^k(\beta) \\ \cup\{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^k(\alpha)z \geq 0, z \leq 0\} \\ \qquad\qquad\qquad \text{if } \bar{v} = \nabla\vartheta^k(\beta). \end{cases}$$

For $\bar{p} = \kappa^j$ with $j \in \{1, 2, ..., k\}$, there are two cases:

(a) $j \in M$. Then

$$N((\kappa^j, \bar{v}); \text{gph}\partial\varphi) = \begin{cases} \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \text{ if } \nabla\vartheta^{j-1}(\kappa^j) < \bar{v} < \nabla\vartheta^j(\kappa^j), \\ \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \cup \mathcal{A}^{j-1}(\kappa^j) \\ \cup\{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^{j-1}(\kappa^j)z \geq 0, z \leq 0\} \\ \qquad\qquad\qquad\qquad \text{if } \bar{v} = \nabla\vartheta^{j-1}(\kappa^j), \\ \{(w, z) \in \mathbb{R}^2 \mid z = 0\} \cup \mathcal{A}^j(\kappa^j) \\ \cup\{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^j(\kappa^j)z \leq 0, z \geq 0\} \\ \qquad\qquad\qquad\qquad \text{if } \bar{v} = \nabla\vartheta^j(\kappa^j), \end{cases}$$

provided that $\nabla\vartheta^{j-1}(\kappa^j) < \nabla\vartheta^j(\kappa^j)$, and

$$N((\kappa^j, \bar{v}); \text{gph}\partial\varphi) = \mathcal{A}^{j-1}(\kappa^j) \cup \mathcal{A}^j(\kappa^j)$$
$$\cup \left\{(w, z) \in \mathbb{R}^2 \mid -\nabla^2\vartheta^{j-1}(\kappa^j)z \leq w \leq -\nabla^2\vartheta^j(\kappa^j)z\right\},$$

provided that $\bar{v} = \nabla\vartheta^{j-1}(\kappa^j) = \nabla\vartheta^j(\kappa^j)$.

(b) $j \notin M$. Then $\bar{v}$ cannot lie between $\nabla\vartheta^{j-1}(\kappa^j)$ and $\nabla\vartheta^j(\kappa^j)$, and hence one has

$$N((\kappa^j, \bar{v}); \text{gph}\partial\varphi)$$
$$= \begin{cases} \{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^{j-1}(\kappa^j)z \geq 0\} \text{ if } \bar{v} = \nabla\vartheta^{j-1}(\kappa^j), \\ \{(w, z) \in \mathbb{R}^2 \mid w + \nabla^2\vartheta^j(\kappa^j)z \leq 0\} \quad \text{if } \bar{v} = \nabla\vartheta^j(\kappa^j). \end{cases} \tag{4.17}$$

We now focus on the class of separable extended-real-valued functions of many variables $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by

$$\varphi(x) = \sum_{i=1}^{n} \varphi_i(x_i), \tag{4.18}$$

where each $\varphi_i : \mathbb{R} \to \bar{\mathbb{R}}$ has the structure (Eq. 4.16) and satisfies all the assumptions imposed in (i) and (ii). This class of functions is important for the studying mathematical programs with equilibrium constraints which frequently arise in modeling some mechanical equilibria [17].

Theorem 4.10 *Let $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ be defined as in Eq. 4.18. If the PSD property (Eq. 3.1) is satisfied, then φ is convex.*

Proof Case 1: $n = 1$. Then φ is defined as in Eq. 4.16. Since φ is continuous on $[\alpha, \beta] \cap \mathbb{R}$ and $\varphi(x) = +\infty$ for $x \notin [\alpha, \beta]$, it suffices to prove that φ is convex on (α, β). We will show that $M = \{1, 2, ..., k\}$. To the contrary suppose that there exists an index $j \in \{1, 2, ..., k\} \backslash M$. Then Eq. 4.17 holds. Thus, for $(x, y) = (\kappa^j, \nabla\vartheta^{j-1}(\kappa^j))$,

$$z \in \partial^2\varphi(x, y)(u) \Longleftrightarrow z - \nabla^2\vartheta^{j-1}(\kappa^j)u \geq 0.$$

Hence for $u = -1$ and $z = |\nabla^2\vartheta^{j-1}(\kappa^j)| + 1$, we have $z \in \partial^2\varphi(x, y)(u)$ and $zu < 0$. This contradicts Eq. 3.1. Hence $M = \{1, 2, ..., k\}$. Since φ is twice continuously differentiable at each $x \in (\alpha, \beta)\backslash\{\kappa^1, \kappa^2, ..., \kappa^k\}$, $\partial^2\varphi(x, y)(u) = \{\nabla^2\varphi(x)^*(u)\}$, where

$y = \nabla\varphi(x)$. By Eq. 3.1, $\nabla^2\varphi(x) \geq 0$. It follows that $\partial\varphi(\cdot) = \{\nabla\varphi(\cdot)\}$ is monotone on each one of the intervals (α, κ^1), (κ^j, κ^{j+1}) $(j = 1, 2, ..., k-1)$, and (κ^k, β). Besides,

$$\partial\varphi(x) = \begin{cases} \{\nabla\vartheta^0(x)\} & \text{if } x \in (\alpha, \kappa^1), \\ \{\nabla\vartheta^j(x)\} & \text{if } x \in (\kappa^j, \kappa^{j+1}) \ (j = 1, 2, ..., k-1), \\ \{\nabla\vartheta^k(x)\} & \text{if } x \in (\kappa^k, \beta), \\ [\nabla\vartheta^{j-1}(\kappa^j), \nabla\vartheta^j(\kappa^j)] & \text{if } x = \kappa^j \ (j \in M = \{1, 2, ..., k\}). \end{cases}$$

Hence $\partial\varphi(\cdot)$ is monotone on (α, β) and, therefore, φ is convex on (α, β).

Case 2: $n > 1$. For any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ with $y \in \partial\varphi(x)$ and $u \in \mathbb{R}^n$,

$$\partial^2\varphi(x, y)(u) = \{\omega \in \mathbb{R}^n \mid (\omega_i, -u_i) \in N((x_i, y_i); \text{gph}\partial\varphi_i), \ i = 1, 2, ..., n\};$$

see [17, Theorem 4.3]. Hence Eq. 3.1 yields $z_i u_i \geq 0$ for all $u_i \in \mathbb{R}$, $z_i \in \partial^2\varphi_i(x_i, y_i)(u_i)$ with $(x_i, y_i) \in \text{gph}\partial\varphi_i$ $(i = 1, 2, ..., n)$. According to the analysis already given in Case 1, φ_i is convex for $i = 1, 2, ..., n$. Hence φ is convex. □

5 Characterizations of Strong Convexity

Strong convexity of real-valued functions in a neighborhood of a given point is of a frequent use in formulating sufficient optimality conditions and stability criteria for extremum problems, in theory of optimization algorithms. We are going to derive from the above results some criteria for strong convexity of real-valued functions.

A function $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be *strongly convex* on a convex subset $\Omega \subset \text{dom}\varphi$ if there exists a constant $\rho > 0$ such that the inequality

$$\varphi((1-t)x + ty) \leq (1-t)\varphi(x) + t\varphi(y) - \rho t(1-t)\|x - y\|^2$$

holds for any $x, y \in \Omega$ and $t \in (0, 1)$. It is well known (see, e.g., [24, Lemma 1, p. 184]) that the last condition is fulfilled if and only if the function

$$\widetilde{\varphi}(x) := \varphi(x) - \rho\|x\|^2 \tag{5.1}$$

is convex on Ω.

Using second-order subdifferential mappings, we can formulate necessary conditions for strong convexity of a real function as follows.

Theorem 5.1 *Let $\varphi : \mathbb{R}^n \to \bar{\mathbb{R}}$ be proper lower semicontinuous. If φ is strongly convex on $\mathbb{R}^n$ with the constant $\rho > 0$, then for any $(x, y) \in \text{gph}\partial\varphi$ the second-order subdifferential mappings $\partial^2\varphi(x, y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $\widehat{\partial}^2\varphi(x, y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfy the conditions*

$$\langle z, u\rangle \geq 2\rho\|u\|^2 \quad \textit{for all } u \in \mathbb{R}^n \textit{ and } z \in \partial^2\varphi(x, y)(u) \textit{ with } (x, y) \in \text{gph}\partial\varphi \tag{5.2}$$

and

$$\langle z, u\rangle \geq 2\rho\|u\|^2 \quad \textit{for all } u \in \mathbb{R}^n, z \in \widehat{\partial}^2\varphi(x, y)(u) \textit{ with } (x, y) \in \text{gph}\widehat{\partial}\varphi. \tag{5.3}$$

Proof Since Eq. 5.2 implies Eq. 5.3, it suffices to prove Eq. 5.2. By our assumption, the function $\widetilde{\varphi}$ given by Eq. 5.1 is convex on $\mathbb{R}^n$. Applying the subdifferential sum rule with equality [16, Prop. 1.107(ii)] for the sum $\widetilde{\varphi} = \varphi + \psi$, where $\psi(x) = -\rho\|x\|^2$, we find

$$\partial\widetilde{\varphi}(x) = \partial\varphi(x) - 2\rho x \quad \forall x \in \mathbb{R}^n. \tag{5.4}$$

(Note that, in the case under consideration, $\partial\varphi(x)$ consists of those $\xi \in \mathbb{R}^n$ with the property that $\langle \xi, v - x\rangle \le \varphi(v) - \varphi(x)$ for all $v \in \mathbb{R}^n$.) Now, setting $F(x) = \partial\varphi(x)$, $f(x) = -2\rho x$, and using the coderivative sum rule with equality [16, Prop. 1.62(ii)], we get

$$D^*(F + f)(x, y - 2\rho x)(u) = D^*F(x, y)(u) - 2\rho u$$

for any $x \in \mathbb{R}^n$, $y \in \partial\varphi(x)$, and $u \in \mathbb{R}^n$. The latter in a combination with Eq. 5.4 implies that

$$\partial^2\widetilde{\varphi}(x, y - 2\rho x)(u) = \partial^2\varphi(x, y)(u) - 2\rho u \quad \forall x \in \mathbb{R}^n,\ \forall y \in \partial\varphi(x),\ \forall u \in \mathbb{R}^n. \tag{5.5}$$

According to Theorem 3.2, from the convexity of $\widetilde{\varphi}$ it follows that the second-order subdifferential mapping $\partial^2\widetilde{\varphi}(\cdot)$ is PSD. Hence, by Eq. 5.5 we obtain $\langle z - 2\rho u, u\rangle \ge 0$ for any $z \in \partial^2\varphi(x, y)(u)$; thus Eq. 5.2 holds for every $(x, y) \in \text{gph}\partial\varphi$. □

The next proposition describes sufficient conditions for the strong convexity of some classes of real functions.

Theorem 5.2 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a function and $\rho > 0$ a given number. The following assertions are valid:*

(i) *If φ is a $C^{1,1}$ function and if Eq. 5.3 is fulfilled for any $(x, y) \in \text{gph}\partial\varphi$, then φ is strongly convex on $\mathbb{R}^n$ with the constant $\rho > 0$.*
(ii) *If $n = 1$, φ is a C^1 function, and the condition Eq. 5.3 is fulfilled for any $(x, y) \in \text{gph}\partial\varphi$, then φ is strongly convex on $\mathbb{R}$ with the constant $\rho > 0$.*
(iii) *If $\varphi(x) = \sum_{i=1}^{n} \varphi_i(x_i)$, where each $\varphi_i : \mathbb{R} \to \bar{\mathbb{R}}$ has the structure (Eq. 4.16), and if Eq. 5.2 is fulfilled for any $(x, y) \in \text{gph}\partial\varphi$, then φ is strongly convex on $\mathbb{R}^n$ with the constant $\rho > 0$.*

Proof Define $\widetilde{\varphi}$, F, and f as in the above proof.

(i) By the sum rules with equalities for Fréchet subdifferentials and coderivatives [16, Props. 1.107(i) and 1.62(i)], we can perform some simple calculations as those in the proof of Theorem 5.1 to get

$$\widehat{\partial}^2\widetilde{\varphi}(x, \nabla\varphi(x) - 2\rho x)(u) = \widehat{\partial}^2\varphi(x, \nabla\varphi(x))(u) - 2\rho u \quad \forall x \in \mathbb{R}^n,\ \forall u \in \mathbb{R}^n. \tag{5.6}$$

From Eqs. 5.3 and 5.6 it follows that $\widehat{\partial}^2\widetilde{\varphi}(x, \nabla\varphi(x) - 2\rho x)$ is PSD for any $x \in \mathbb{R}^n$. Thus, the $C^{1,1}$ function $\widetilde{\varphi}$ is convex on $\mathbb{R}^n$ according to Theorem 4.2. Then φ is strongly convex on $\mathbb{R}^n$ with the constant $\rho > 0$.
(ii) Since the sum rules with equalities for Fréchet subdifferentials and coderivatives [16, Props. 1.107(i) and 1.62(i)] are applicable to the present C^1 setting, Eq. 5.6 is valid. Hence the C^1 function $\widetilde{\varphi} : \mathbb{R} \to \mathbb{R}$ is convex by Theorem 4.4, and the desired conclusion follows.

(iii) Applying Eq. 5.5 and the assumptions made, we can assert that $\partial^2\widetilde{\varphi}(\cdot)$ is PSD. Since

$$\widetilde{\varphi}(x) = \sum_{i=1}^{n} \left(\varphi_i(x_i) - \rho x_i^2\right)$$

has the decomposable structure required by Theorem 4.10, it follows that $\widetilde{\varphi}$ is convex on $\mathbb{R}^n$; hence φ is strongly convex on $\mathbb{R}^n$ with the constant $\rho > 0$. □

6 Open Questions

We have obtained several necessary and sufficient conditions for the convexity of extended-real-valued functions on finite-dimensional Euclidean spaces.

The following questions remain open, thus require further investigations:

1. Is it true that, for any C^1 function $\varphi : \mathbb{R}^n \to \mathbb{R}$ with $n \geq 2$, condition 3.2 implies the convexity of φ?
2. Is it true that, for any locally Lipschitz function $\varphi : \mathbb{R}^n \to \mathbb{R}$, condition 3.1 implies the convexity of φ?
3. Is it true that, for any continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$, condition 3.1 implies the convexity of φ?
4. How to extend the obtained results to the infinite-dimensional Hilbert space setting and, furthermore, to an infinite-dimensional Banach space setting?

References

1. Bednarik, D., Pastor, K.: On characterizations of convexity for regularly locally Lipschitz functions. Nonlinear Anal. **57**, 85–97 (2004)
2. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
3. Cohn, D.L.: Measure Theory. Birkhäuser, Boston (1980)
4. Cominetti, R., Correa, R.: A generalized second-order derivative in nonsmooth optimization. SIAM J. Control Optim. **28**, 789–809 (1990)
5. Dontchev, A.L., Rockafellar, R.T.: Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **6**, 1087–1105 (1996)
6. Ginchev, I., Ivanov, V.I.: Second-order characterizations of convex and pseudoconvex functions. J. Appl. Anal. **9**, 261–273 (2003)
7. Hadjisavvas, N., Komlósi, S., Schaible, S. (eds.): Handbook of Generalized Convexity and Generalized Monotonicity. Springer, New York (2005)
8. Hiriart-Urruty, J.-B., Strodiot, J.-J., Nguyen, V.H.: Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data. Appl. Math. Optim. **11**, 43–56 (1984)
9. Jeyakumar, V., Yang, X.Q.: Approximate generalized Hessians and Taylor's expansions for continuously Gâteaux differentiable functions. Nonlinear Anal. **36**, 353–368 (1999)
10. Huang, L.R., Ng, K.F.: On lower bounds of the second-order directional derivatives of Ben-Tal, Zowe, and Chaney. Math. Oper. Res. **22**, 747–753 (1997)
11. Lee, G.M., Kim, D.S., Lee, B.S., Yen, N.D.: Vector variational inequality as a tool for studying vector optimization problems. Nonlinear Anal. **34**, 745–765 (1998)
12. Lee, G.M., Tam, N.N., Yen, N.D.: Quadratic Programming and Affine Variational Inequalities: a Qualitative Study. Springer, New York (2005)
13. Levy, A.B., Poliquin, R.A., Rockafellar, R.T.: Stability of locally optimal solutions. SIAM J. Optim. **10**, 580–604 (2000)

14. Mordukhovich, B.S.: Sensitivity analysis in nonsmooth optimization. In: Field , D.A., Komkov, V. (eds.) Theoretical Aspects of Industrial Design. SIAM Proc. in Applied Mathematics, vol. 58, pp. 32–42. SIAM, Philadelphia (1992)
15. Mordukhovich, B.S.: Generalized differential calculus for nonsmooth and set-valued mappings. J. Math. Anal. Appl. **183**, 250–288 (1994)
16. Mordukhovich, B.S.: Variational analysis and generalized differentiation. In: Basic Theory, vol. I. Applications, vol. II. Springer, Berlin (2006)
17. Mordukhovich, B.S., Outrata, J.V.: On second-order subdifferentials and their applications. SIAM J. Optim. **12**, 139–169 (2001)
18. Phelps, R.R.: Convex functions, monotone operators and differentiability. In: Lecture Notes in Math, vol. 1364. Springer, Berlin (1993)
19. Poliquin, R.A., Rockafellar, R.T.: Tilt stability of a local minimum. SIAM J. Optim. **8**, 287–299 (1998)
20. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
21. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
22. Rudin, W.: Real and Complex Analysis, 2nd edn. McGraw-Hill, New York (1974)
23. Segal, I.E., Kunze, R.A.: Integrals and Operators. Springer, Berlin (1978)
24. Vasilev, F.P.: Numerical Methods for Solving Extremal Problems, 2nd edn. Nauka, Moscow (In Russian) (1988)
25. Yang, X.Q.: Generalized second-order characterizations of convex functions. J. Optim. Theory Appl. **82**, 173–180 (1994)
26. Yang, X.Q., Jeyakumar, V.: Generalized second-order directional derivatives and optimization with $C^{1,1}$ functions. Optimization **26**, 165–185 (1992)
27. Yen, N.D.: Hölder continuity of solutions to a parametric variational inequality. Appl. Math. Optim. **31**, 245–255 (1995)
28. Yen, N.D.: Lipschitz continuity of solutions of variational inequalities with a parametric polyhedral constraint. Math. Oper. Res. **20**, 695–708 (1995)