TWICE EPI-DIFFERENTIABILITY OF EXTENDED-REAL-VALUED FUNCTIONS WITH APPLICATIONS IN COMPOSITE OPTIMIZATION

ASHKAN MOHAMMADI[1] and M. EBRAHIM SARABI[2]

Abstract. The paper is devoted to the study of the twice epi-differentiablity of extended-real-valued functions, with an emphasis on functions satisfying a certain composite representation. This will be conducted under parabolic regularity, a second-order regularity condition that was recently utilized in [14] for second-order variational analysis of constraint systems. Besides justifying the twice epi-differentiablity of composite functions, we obtain precise formulas for their second subderivatives under the metric subregularity constraint qualification. The latter allows us to derive second-order optimality conditions for a large class of composite optimization problems.

Key words. Variational analysis, twice epi-differentiability, parabolic regularity, composite optimization, second-order optimality conditions

Mathematics Subject Classification (2000) 49J53, 49J52, 90C31

1 Introduction

This paper aims to provide a systematic study of the twice epi-differentiability of extend-real-valued functions in finite dimensional spaces. In particular, we pay special attention to the composite optimization problem

$$\text{minimize}\ \ \varphi(x) + g(F(x)) \quad \text{over all}\ x \in \mathbb{X}, \tag{1.1}$$

where $\varphi : \mathbb{X} \to \mathbb{R}$ and $F : \mathbb{X} \to \mathbb{Y}$ are twice differentiable and $g : \mathbb{Y} \to \overline{\mathbb{R}} := (-\infty, +\infty]$ is a lower semicontinuous (l.s.c.) convex function and where $\mathbb{X}$ and $\mathbb{Y}$ are two finite dimensional spaces, and verify the twice epi-differentiability of the objective function in (1.1) under verifiable assumptions. The composite optimization problem (1.1) encompasses major classes of constrained and composite optimization problems including classical nonlinear programming problems, second-order cone and semidefinite programming problems, eigenvalue optimizations problems [24], and fully amenable composite optimization problems [19], see Example 4.7 for more detail. Consequently, the composite problem (1.1) provides a unified framework to study second-order variational properties, including the twice epi-differentiability and second-order optimality conditions, of the aforementioned optimization problems. As argued below, the twice epi-differentiability carries vital second-order information for extend-real-valued functions and therefore plays an important role in modern second-order variational analysis.

A lack of an appropriate second-order generalized derivative for nonconvex extended-real-valued functions was the main driving force for Rockafellar to introduce in [17] the concept of the twice epi-differentiability for such functions. Later, in his landmark paper [19], Rockafellar justified this property for an important class of functions, called *fully amenable*, that includes nonlinear programming problems but does not go far enough to cover other major classes of constrained and composite optimization problems. Rockafellar's results were extended in [7,10] for composite functions appearing in (1.1). However, these extensions were achieved under a

---

[1]Department of Mathematics, Wayne State University, Detroit, MI 48202, USA (ashkan.mohammadi@wayne.edu).

[2]Department of Mathematics, Miami University, Oxford, OH 45065, USA (sarabim@miamioh.edu).

restrictive assumption on the second subderivative, which does not hold for constrained optimization problems. Nor does this condition hold for other major composite functions related to eigenvalue optimization problems; see [24, Theorem 1.2] for more detail. Levy in [11] obtained upper and lower estimates for the second subderivative of the composite function from (1.1), but fell short of establishing the twice epi-differentiability for this framework.

The authors and Mordukhovich observed recently in [14] that a second-order regularity, called *parabolic regularity* (see Definition 3.1), can play a major role toward the establishment of the twice epi-differentiability for constraint systems, namely when the outer function g in (1.1) is the indicator function of a closed convex set. This vastly alleviated the difficulty that was often appeared in the justification of the twice epi-differentiability for the latter framework and opened the door for crucial applications of this concept in theoretical and numerical aspects of optimization. Among these applications, we can list the following:

- the calculation of proto-derivatives of subgradient mappings via the connection between the second subderivative of a function and the proto-derivative of its subgradient mapping (see equation (3.21));
- the calculation of the second subderivative of the augmented Lagrangian function associated with the composite problem (1.1), which allows us to characterize the second-order growth condition for the augmented Lagrangian problem (cf. [14, Theorems 8.3 & 8.4]);
- the validity of the derivative-coderivative inclusion (cf. [21, Theorem 13.57]), which has important consequences in parametric optimization; see [15, Theorem 5.6] for a recent application in the convergence analysis of the sequential quadratic programming (SQP) method for constrained optimization problems.

In this paper, we continue the path, initiated in [14] for constraint systems, and show that the twice epi-differentiability of the objective function in (1.1) can be guaranteed under parabolic regularity. To achieve this goal, we demand that the outer function g from (1.1) be locally Lipschitz continuous relative to its domain; see the next section for the precise definition of this concept. Shapiro in [22] used a similar condition but in addition assumed that this function is finite-valued. The latter does bring certain restrictions for (1.1) by excluding constrained problems as well as piecewise linear-quadratic composite problems. As shown in Example 4.7, major classes of constrained and composite optimization problems satisfy this Lipschitzian condition. However, some composite problems such as the spectral abcissa minimization (cf. [4]), namely the problem of minimizing the largest real parts of eigenvalues, can not be covered by (1.1).

The rest of the paper is organized as follows. Section 2 recalls important notions of variational analysis that are used throughout this paper. Section 3 begins with the definition of parabolic regularity of extended-real-valued functions. Then we justify that parabolic regularity amounts to a certain duality relationship between the second subderivative and parabolic subderivative. Employing this, we show that the twice epi-differentiability of extended-real-valued functions can be guaranteed if they are parabolically regular and parabolic epi-differentiable. Section 4 is devoted to important second-order variational properties of parabolic subderivatives. In particular, we establish a chain rule for parabolic subderivatives of composite functions in (1.1) under the metric subregularity constraint qualification. In Section 5, we establish chain rules for the parabolic regularity and for the second subderivative of composite functions, and consequently establish their twice epi-differentiability. Section 6 deals with important applications of our results in second-order optimality conditions for the composite optimization problem (1.1). We close the paper by achieving a characterization of the strong metric subregularity of the subgradient mapping of the objective function in (1.1) via the second-order sufficient condition for this problem.

In what follows, $\mathbb{X}$ and $\mathbb{Y}$ are finite-dimensional Hilbert spaces equipped with a scalar product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. By $\mathbb{B}$ we denote the closed unit ball in the space in question

and by $\mathbb{B}_r(x) := x + r\mathbb{B}$ the closed ball centered at x with radius $r > 0$. For any set C in $\mathbb{X}$, its indicator function is defined by $\delta_C(x) = 0$ for $x \in C$ and $\delta_C(x) = \infty$ otherwise. We denote by $d(x, C)$ the distance between $x \in \mathbb{X}$ and a set C. For $v \in \mathbb{X}$, the subspace $\{w \in \mathbb{X} | \langle w, v \rangle = 0\}$ is denoted by $\{v\}^\perp$. We write $x(t) = o(t)$ with $x(t) \in \mathbb{X}$ and $t > 0$ to mean that $\|x(t)\|/t$ goes to 0 as $t \downarrow 0$. Finally, we denote by $\mathbb{R}_+$ (respectively, $\mathbb{R}_-$) the set of non-negative (respectively, non-positive) real numbers.

2 Preliminary Definitions in Variational Analysis

In this section we first briefly review basic constructions of variational analysis and generalized differentiation employed in the paper; see [12, 21] for more detail. A family of sets C_t in $\mathbb{X}$ for $t > 0$ converges to a set $C \subset \mathbb{X}$ if C is closed and

$$\lim_{t \downarrow 0} d(w, C_t) = d(w, C) \quad \text{for all} \ \ w \in \mathbb{X}.$$

Given a nonempty set $C \subset \mathbb{X}$ with $\bar{x} \in C$, the tangent cone $T_C(\bar{x})$ to C at $\bar{x}$ is defined by

$$T_C(\bar{x}) = \{w \in \mathbb{X} | \ \exists t_k \downarrow 0, \ w_k \to w \ \ \text{as} \ \ k \to \infty \ \ \text{with} \ \ \bar{x} + t_k w_k \in C\}.$$

We say a tangent vector $w \in T_C(\bar{x})$ is *derivable* if there exist a constant $\varepsilon > 0$ and an arc $\xi : [0, \varepsilon] \to C$ such that $\xi(0) = \bar{x}$ and $\xi'_+(0) = w$, where ξ'_+ signifies the right derivative of ξ at 0, defined by

$$\xi'_+(0) := \lim_{t \downarrow 0} \frac{\xi(t) - \xi(0)}{t}.$$

The set C is called geometrically derivable at $\bar{x}$ if every tangent vector w to C at $\bar{x}$ is derivable. The geometric derivability of C at $\bar{x}$ can be equivalently described by the sets $[C - \bar{x}]/t$ converging to $T_C(\bar{x})$ as $t \downarrow 0$. Convex sets are important examples of geometrically derivable sets. The second-order tangent set to C at $\bar{x}$ for a tangent vector $w \in T_C(\bar{x})$ is given by

$$T_C^2(\bar{x}, w) = \{u \in \mathbb{X} | \ \exists t_k \downarrow 0, \ u_k \to u \ \ \text{as} \ \ k \to \infty \ \ \text{with} \ \ \bar{x} + t_k w + \frac{1}{2} t_k^2 u_k \in C\}.$$

A set C is said to be parabolically derivable at $\bar{x}$ for w if $T_C^2(\bar{x}, w)$ is nonempty and for each $u \in T_C^2(\bar{x}, w)$ there are $\varepsilon > 0$ and an are $\xi : [0, \varepsilon] \to C$ with $\xi(0) = \bar{x}$, $\xi'_+(0) = w$, and $\xi''_+(0) = u$, where

$$\xi''_+(0) := \lim_{t \downarrow 0} \frac{\xi(t) - \xi(0) - t\xi'_+(0)}{\frac{1}{2} t^2}.$$

It is well known that if $C \subset \mathbb{X}$ is convex and parabolically derivable at $\bar{x}$ for w, then the second-order tangent set $T_C^2(\bar{x}, w)$ is a nonempty convex set in $\mathbb{X}$ (cf. [2, page 163]). Given the function $f : \mathbb{X} \to \overline{\mathbb{R}} := (-\infty, \infty]$, its domain and epigraph are defined, respectively, by

$$\text{dom} f = \{x \in \mathbb{X} | \ f(x) < \infty\} \quad \text{and} \quad \text{epi} f = \{(x, \alpha) \in \mathbb{X} \times \mathbb{R} | \ f(x) \le \alpha\}.$$

The regular subdifferential of f at $\bar{x} \in \text{dom} f$ is defined by

$$\widehat{\partial} f(\bar{x}) = \{v \in \mathbb{X} | \ \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \ge 0\}.$$

The subdifferential of f at $\bar{x}$ is given by

$$\partial f(\bar{x}) = \{v \in \mathbb{X} | \ \exists x_k \xrightarrow{f} \bar{x}, \ v_k \to v \ \text{with} \ v_k \in \widehat{\partial} f(x_k)\},$$

3

where $x_k \xrightarrow{f} \bar{x}$ stands for $x_k \to \bar{x}$ and $f(x_k) \to f(\bar{x})$. We say that $v \in \mathbb{X}$ is a proximal subgradient of f at $\bar{x}$ if there exists $r \in \mathbb{R}_+$ and a neighborhood U of $\bar{x}$ such that for all $x \in U$ we have

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle - \frac{r}{2} \|x - \bar{x}\|^2. \tag{2.1}$$

The set of all such v is called the proximal subdifferential of f at $\bar{x}$ and is denoted by $\partial^p f(\bar{x})$. By definitions, it is not hard to obtain the inclusions $\partial^p f(\bar{x}) \subset \widehat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$. Given a nonempty set $C \subset \mathbb{X}$, the proximal and regular normal cones to C at $\bar{x} \in C$ are defined, respectively, by

$$N_C^p(\bar{x}) := \partial^p \delta_C(\bar{x}) \quad \text{and} \quad \widehat{N}_C(\bar{x}) := \widehat{\partial} \delta_C(\bar{x}).$$

Similarly, we define the (limiting/Mordukhovich) normal cone of C at $\bar{x}$ by $N_C(\bar{x}) := \partial \delta_C(\bar{x})$. Consider a set-valued mapping $S : \mathbb{X} \rightrightarrows \mathbb{Y}$ with its domain and graph defined, respectively, by

$$\text{dom}\, S = \big\{ x \in \mathbb{X} |\ S(x) \neq \emptyset \big\} \quad \text{and} \quad \text{gph}\, S = \big\{ (x,y) \in \mathbb{X} \times \mathbb{Y} |\ y \in S(x) \big\}.$$

The graphical derivative of S at $(\bar{x}, \bar{y}) \in \text{gph}\, S$ is defined by

$$DS(\bar{x}, \bar{y})(w) = \big\{ v \in \mathbb{Y} |\ (w, v) \in T_{\text{gph}\, S}(\bar{x}, \bar{y}) \big\}, \quad w \in \mathbb{X}.$$

Recall that a set-valued mapping $S \colon \mathbb{X} \rightrightarrows \mathbb{Y}$ is metrically regular around $(\bar{x}, \bar{y}) \in \text{gph}\, S$ if there are constants $\kappa \in \mathbb{R}_+$ and $\varepsilon > 0$ such that the distance estimate

$$d\big(x, S^{-1}(y)\big) \leq \kappa\, d\big(y, S(x)\big) \ \text{ for all } \ (x, y) \in \mathbb{B}_\varepsilon(\bar{x}, \bar{y})$$

holds. When $y = \bar{y}$ in the above estimate, the mapping S is called metrically subregular at $(\bar{x}, \bar{y})$. The set-valued mapping S is called strongly metrically subregular at $(\bar{x}, \bar{y})$ if there are a constant $\kappa \in \mathbb{R}_+$ and a neighborhood U of $\bar{x}$ such that the estimate

$$\|x - \bar{x}\| \leq \kappa\, d(\bar{y}, S(x)) \ \text{ for all } \ x \in U$$

holds. It is known (cf. [8, Theorem 4E.1]) that the set-valued mapping S is strongly metrically subregular at $(\bar{x}, \bar{y})$ if and only if we have

$$0 \in DS(\bar{x}, \bar{y})(w) \implies w = 0. \tag{2.2}$$

Given a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ and a point $\bar{x}$ with $f(\bar{x})$ finite, the subderivative function $\mathrm{d}f(\bar{x}) \colon \mathbb{R}^n \to [-\infty, \infty]$ is defined by

$$\mathrm{d}f(\bar{x})(\bar{w}) = \liminf_{\substack{t \downarrow 0 \\ w \to \bar{w}}} \frac{f(\bar{x} + tw) - f(\bar{x})}{t}.$$

Define the parametric family of second-order difference quotients for f at $\bar{x}$ for $\bar{v} \in \mathbb{X}$ by

$$\Delta_t^2 f(\bar{x}, \bar{v})(w) = \frac{f(\bar{x} + tw) - f(\bar{x}) - t \langle \bar{v},\, w \rangle}{\frac{1}{2} t^2} \qquad \text{with } \ w \in \mathbb{X}, \ \ t > 0.$$

If $f(\bar{x})$ is finite, then the second subderivative of f at $\bar{x}$ for $\bar{v}$ is given by

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \liminf_{\substack{t \downarrow 0 \\ w' \to w}} \Delta_t^2 f(\bar{x}, \bar{v})(w'), \ \ w \in \mathbb{X}.$$

Below, we collect some important properties of the second subderivative that are used throughout this paper. Parts (i) and (ii) were taken from [21, Proposition 13.5] and part (iii) was recently observed in [13, Theorem 4.1(i)].

Proposition 2.1 (properties of second subderivative). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ and $(\bar{x}, \bar{v}) \in \mathbb{X} \times \mathbb{X}$ with $f(\bar{x})$ finite. Then the following conditions hold:*

   (i) *the second subderivative $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a lower semicontinuous (l.s.c.) function;*

   (ii) *if $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function, meaning that $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) > -\infty$ for all $w \in \mathbb{X}$ and its effective domain, defined by*

$$\operatorname{dom} \mathrm{d}^2 f(\bar{x}, \bar{v}) = \left\{ w \in \mathbb{X} \mid \mathrm{d}^2 f(\bar{x}, \bar{v})(w) < \infty \right\},$$

   is nonempty, then we always have the inclusion

$$\operatorname{dom} \mathrm{d}^2 f(\bar{x}, \bar{v}) \subset \left\{ w \in \mathbb{X} \mid \mathrm{d} f(\bar{x})(w) = \langle \bar{v}, w \rangle \right\};$$

   (iii) *if $\bar{v} \in \partial^p f(\bar{x})$, then for any $w \in \mathbb{X}$ we have $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) \geq -r \|w\|^2$, where $r \in \mathbb{R}_+$ is taken from (2.1). In particular, $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function.*

Following [21, Definition 13.6], a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is said to be twice epi-differentiable at $\bar{x}$ for $\bar{v} \in \mathbb{X}$, with $f(\bar{x})$ finite, if the sets $\operatorname{epi} \Delta_t^2 f(\bar{x}, \bar{v})$ converge to $\operatorname{epi} \mathrm{d}^2 f(\bar{x}, \bar{v})$ as $t \downarrow 0$. The latter means by [21, Proposition 7.2] that for every sequence $t_k \downarrow 0$ and every $w \in \mathbb{X}$, there exists a sequence $w_k \to w$ such that

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \lim_{k \to \infty} \Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k). \tag{2.3}$$

We say that a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is Lipschitz continuous around $\bar{x}$ *relative* to $C \subset \operatorname{dom} f$ with constant $\ell \in \mathbb{R}_+$ if $\bar{x} \in C$ and there exists a neighborhood U of $\bar{x}$ such that

$$|f(x) - f(y)| \leq \ell \|x - y\| \quad \text{for all} \ \ x, y \in U \cap C.$$

Such a function is called *locally* Lipschitz continuous relative to C if for every $\bar{x} \in C$, it is Lipschitz continuous around $\bar{x}$ relative to C. Piecewise linear-quadratic functions (not necessarily convex) and an indicator function of a nonempty set are important examples of functions that are locally Lipschitz continuous relative to their domains.

Proposition 2.2 (domain of subderivatives). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ be Lipschitz continuous around $\bar{x}$ relative to its domain with constant $\ell \in \mathbb{R}_+$. Then we have $\operatorname{dom} \mathrm{d} f(\bar{x}) = T_{\operatorname{dom} f}(\bar{x})$. In particular, for every $w \in T_{\operatorname{dom} f}(\bar{x})$, the subderivative $\mathrm{d} f(\bar{x})(w)$ is finite.*

Proof. The inclusion $\operatorname{dom} \mathrm{d} f(\bar{x}) \subset T_{\operatorname{dom} f}(\bar{x})$ results directly from the definition. To prove the opposite inclusion, pick $w \in T_{\operatorname{dom} f}(\bar{x})$. This gives us some sequences $t_k \downarrow 0$ and $w_k \to w$ such that $\bar{x} + t_k w_k \in \operatorname{dom} f$ for all $k \in \mathbb{N}$. Using this and the Lipschitz continuity of f around $\bar{x}$ relative to its domain implies that for all k sufficiently large we have

$$\left| \frac{f(\bar{x} + t_k w_k) - f(\bar{x})}{t_k} \right| \leq \ell \|w_k\|. \tag{2.4}$$

This clearly yields $|\mathrm{d} f(\bar{x})(w)| \leq \ell \|w\|$. Thus $\mathrm{d} f(\bar{x})(w)$ is finite and so $w \in \operatorname{dom} \mathrm{d} f(\bar{x})$. This gives us the inclusion $T_{\operatorname{dom} f}(\bar{x}) \subset \operatorname{dom} \mathrm{d} f(\bar{x})$ and hence completes the proof. $\square$

3 Twice Epi-Differetiability of Parabolically Regular Functions

This section aims to delineate conditions under which the twice epi-differenibility of extend-real-valued functions can be established. To this end, we appeal to an important second-order regularity condition, called parabolic regularity, which was recently exploited in [14] to study a similar property for constraint systems. We begin with the definition of this regularity condition.

Definition 3.1 (parabolic regularity). *A function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is parabolically regular at $\bar{x}$ for $\bar{v} \in \mathbb{X}$ if $f(\bar{x})$ is finite and if for any w such that $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) < \infty$, there exist, among the sequences $t_k \downarrow 0$ and $w_k \to w$ with $\Delta^2_{t_k} f(\bar{x}, \bar{v})(w_k) \to \mathrm{d}^2 f(\bar{x}, \bar{v})(w)$, those with the additional property that*

$$\limsup_{k \to \infty} \frac{\|w_k - w\|}{t_k} < \infty. \tag{3.5}$$

A nonempty set $C \subset \mathbb{X}$ is said to be parabolically regular at $\bar{x}$ for $\bar{v}$ if the indicator function δ_C is parabolically regular at $\bar{x}$ for $\bar{v}$.

Although the notion of parabolic regularity was introduced first in [21, Definition 13.65], its origin goes back to [5, Theorem 4.4], where Chaney observed a duality relationship between his second-order generalized derivative and the parabolic subderivative, defined in [1] by Ben-Tal and Zowe. This duality relationship was derived later by Rockafellar [18, Proposition 3.5] for convex piecewise linear-quadratic functions. As shown in Proposition 3.6 below, the latter duality relationship is equivalent to the concept of parabolic regularity from Definition 3.1 provided that $\bar{v}$, appearing in Definition 3.1, is a proximal subgradient. A different second-order regularity was introduced by Bonnans, Comminetti, and Shapiro [3, Definition 3] for sets, which was later extended in [2, Definition 3.93] for functions. It is not difficult to see that parabolic regularity is implied by the second-order regularity in the sense of [3]; see [2, Proposition 3.103] for a proof of this result. Moreover, the example from [2, page 215] shows that the converse implication may not hold in general.

We showed in [14] that important sets appearing in constrained optimization problems, including polyhedral convex sets, the second-order cone, and the cone of positive semidefinite symmetric matrices, are parabolically regular. Below, we add two important classes of functions for which this property automatically fulfill. We begin first by convex piecewise-linear quadratic functions and then consider eigenvalues functions. While the former was justified in [21, Theorem 13.67], we provide below a different and simpler proof.

Example 3.2 (piecewise linear-quadratic functions). Assume that the function $f : \mathbb{X} \to \overline{\mathbb{R}}$ with $\mathbb{X} = \mathbb{R}^n$ is convex piecewise linear-quadratic. Recall that f is called piecewise linear-quadratic if $\mathrm{dom}\, f = \cup_{i=1}^s C_i$ with $s \in \mathbb{N}$ and C_i being polyhedral convex sets for $i = 1, \ldots, s$, and if f has a representation of the form

$$f(x) = \tfrac{1}{2}\langle A_i x, x \rangle + \langle a_i, x \rangle + \alpha_i \quad \text{for all} \quad x \in C_i,$$

where A_i is an $n \times n$ symmetric matrix, $a_i \in \mathbb{R}^n$, and $\alpha_i \in \mathbb{R}$ for $i = 1, \cdots, s$. It was proven in [21, Propsoition 13.9] that the second subderivative of f at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ can be calculated by

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \begin{cases} \langle A_i w, w \rangle & \text{if } w \in T_{C_i}(\bar{x}) \cap \{\bar{v}_i\}^\perp, \\ \infty & \text{otherwise,} \end{cases} \tag{3.6}$$

where $\bar{v}_i := \bar{v} - A_i \bar{x} - a_i$. To prove the parabolic regularity of f at $\bar{x}$ for $\bar{v}$, pick a vector $w \in \mathbb{R}^n$ with $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) < \infty$. This implies that there is an i with $1 \le i \le s$ such that $w \in T_{C_i}(\bar{x}) \cap \{\bar{v}_i\}^\perp$. Since C_i is a polyhedral convex set, we conclude from [21, Exercise 6.47] that there exists an $\varepsilon > 0$ such that $\bar{x} + tw \in C_i$ for all $t \in [0, \varepsilon]$. Pick a sequence $t_k \downarrow 0$ such

that $t_k \in [0, \varepsilon]$ and let $w_k := w$ for all $k \in \mathbb{N}$. Thus a simple calculation tells us that

$$
\begin{aligned}
\Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) &= \frac{f(\bar{x} + t_k w_k) - f(\bar{x}) - t_k \langle \bar{v}, w_k \rangle}{\frac{1}{2} t_k^2} \\
&= \frac{\frac{1}{2} \langle A_i(\bar{x} + t_k w_k), \bar{x} + t_k w_k \rangle + \langle a_i, \bar{x} + t_k w_k \rangle + \alpha_i - \frac{1}{2} \langle A_i \bar{x}, \bar{x} \rangle - \langle a_i, \bar{x} \rangle - \alpha_i - t_k \langle \bar{v}, w_k \rangle}{\frac{1}{2} t_k^2} \\
&= \langle A_i w, w \rangle + \frac{t_k \langle w_k, \bar{v} - A_i \bar{x} - a_i \rangle}{\frac{1}{2} t_k^2} = \langle A_i w, w \rangle,
\end{aligned}
$$

which in turn implies by (3.6) that $\Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) \to \mathrm{d}^2 f(\bar{x}, \bar{v})(w)$ as $k \to \infty$. Since (3.5) is clearly holds, f is parabolic regular at $\bar{x}$ for $\bar{v}$.

Example 3.3 (eigenvalue functions). Let $\mathbb{X} = \mathbb{S}^n$ be the space of $n \times n$ symmetric real matrices, which is conveniently treated via the inner product

$$
\langle A, B \rangle := \operatorname{tr} AB
$$

with $\operatorname{tr} AB$ standing for the sum of the diagonal entries of AB. For a matrix $A \in \mathbb{S}^n$, we denote by $A^\dagger$ the Moore-Penrose pseudo-inverse of A and by $\operatorname{eig} A = (\lambda_1(A), \ldots, \lambda_n(A))$ the vector of eigenvalues of A in decreasing order with eigenvalues repeated according to their multiplicity. Given $i \in \{1, \ldots, n\}$, denote by $\ell_i(A)$ the number of eigenvalues that are equal to $\lambda_i(A)$ but are ranked before i including $\lambda_i(A)$. This integer allows us to locate $\lambda_i(A)$ in the group of the eigenvalues of A as follows:

$$
\lambda_1(A) \geq \cdots \geq \lambda_{i-\ell_i(A)} > \lambda_{i-\ell_i(A)+1}(A) = \cdots = \lambda_i(A) \geq \cdots \geq \lambda_n(A).
$$

The eigenvalue $\lambda_{i-\ell_i(A)+1}(A)$, ranking first in the group of eigenvalues equal to $\lambda_i(A)$, is called the *leading* eigenvalue. For any $i \in \{1, \ldots, n\}$, define now the function $\alpha_i : \mathbb{S}^n \to \mathbb{R}$ by

$$
\alpha_i(A) = \lambda_{i-\ell_i(A)+1}(A) + \cdots + \lambda_i(A), \quad A \in \mathbb{S}^n. \tag{3.7}
$$

It was proven in [24, Theorem 2.1] that $\widehat{\partial} \alpha_i(A) = \partial \alpha_i(A)$ and that the second subderivative of α_i at A for any $V \in \partial \alpha_i(A)$ is calculated for every $W \in \mathbb{S}^n$ by

$$
\mathrm{d}^2 \alpha_i(A, V)(W) = \begin{cases} 2 \langle V, W(\lambda_i(A) I_n - A)^\dagger W \rangle & \text{if } \mathrm{d}\alpha_i(A)(W) = \langle X, W \rangle, \\ \infty & \text{otherwise}, \end{cases} \tag{3.8}
$$

where I_n stands for the $n \times n$ identity matrix. Moreover, for any $W \in \mathbb{S}^n$ with $\mathrm{d}^2 \alpha_i(A, H)(W) < \infty$ and any sequence $t_k \downarrow 0$, the proof of [24, Theorem 2.1] confirms that

$$
\Delta_{t_k}^2 \alpha_i(A, V)(W_k) \to \mathrm{d}^2 \alpha_i(A, V)(W) \quad \text{with } W_k := W - t_k W(\lambda_i(A) I_n - A)^\dagger W.
$$

This readily verifies (3.5) and thus the functions α_i, $i \in \{1, \ldots, n\}$, are parabolically regular at A for any $V \in \partial \alpha_i(A)$. In particular, for $i = 1$, the function α_i from (3.7) boils down to the maximum eigenvalue function of a matrix, namely

$$
\lambda_{\max}(A) := \alpha_1(A) = \lambda_1(A), \quad A \in \mathbb{S}^n. \tag{3.9}
$$

So the maximum eigenvalue function $\lambda_{\max}$ is parabolically regular at A for any $V \in \partial \lambda_{\max}(A)$. This can be said for any leading eigenvalue $\lambda_{i-\ell_i(A)+1}(A)$ since we have $\alpha_i(B) = \lambda_{i-\ell_i(A)+1}(B)$ for every matrix $B \in \mathbb{S}^n$ sufficiently close to A. Another important function related to the

eigenvalues of a matrix $A \in \mathbb{S}^n$ is the sum of the first i components of eig A with $i \in \{1, \dots, n\}$, namely

$$\sigma_i(A) = \lambda_1(A) + \cdots + \lambda_i(A). \tag{3.10}$$

It is well-known that the functions σ_i are convex (cf. [21, Exercise 2.54]). Moreover, we have $\sigma_i(A) = \alpha_i(A) + \sigma_{i-\ell_i(A)}(A)$. It follows from [24, Proposition 1.3] that $\sigma_{i-\ell_i(A)}$ is twice continuously differentiable ($\mathcal{C}^2$-smooth) on $\mathbb{S}^n$. This together with the parabolic regularity of α_i ensures that σ_i are parabolically regular at A for any $V \in \partial\sigma_i(A)$.

To proceed further in this section, we require the concept of the parabolic subderivative, introduced by Ben-Tal and Zowe in [1]. Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ and let $\bar{x} \in \mathrm{dom}\, f$ and $w \in \mathbb{X}$ with $\mathrm{d}f(\bar{x})(w)$ finite. The *parabolic subderivative* of f at $\bar{x}$ for w with respect to z is defined by

$$\mathrm{d}^2 f(\bar{x})(w \,|\, z) := \liminf_{\substack{t \downarrow 0 \\ z' \to z}} \frac{f(\bar{x} + tw + \frac{1}{2}t^2 z') - f(\bar{x}) - t\mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t^2}.$$

Recall from [21, Definition 13.59] that f is called *parabolically epi-differentiable* at $\bar{x}$ for w if

$$\mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) = \big\{ z \in \mathbb{X} \,|\, \mathrm{d}^2 f(\bar{x})(w \,|\, z) < \infty \big\} \neq \emptyset,$$

and for every $z \in \mathbb{X}$ and every sequence $t_k \downarrow 0$ there exists a sequences $z_k \to z$ such that

$$\mathrm{d}^2 f(\bar{x})(w \,|\, z) = \lim_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2}. \tag{3.11}$$

The main interest in parabolic subderivatives in this paper lies in its nontrivial connection with second subderivatives. Indeed, it was shown in [21, Proposition 13.64] that if the function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is finite at $\bar{x}$, then for any pair $(\bar{v}, w) \in \mathbb{X} \times \mathbb{X}$ with $\mathrm{d}f(\bar{x})(w) = \langle w, \bar{v} \rangle$ we always have

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) \leq \inf_{z \in \mathbb{X}} \big\{ \mathrm{d}^2 f(\bar{x})(w \,|\, z) - \langle z, \bar{v} \rangle \big\}. \tag{3.12}$$

As observed below, equality in this estimate amounts to the parabolic regularity of f at $\bar{x}$ for $\bar{v}$. To proceed, let $f : \mathbb{X} \to \overline{\mathbb{R}}$ and pick $(\bar{x}, \bar{v}) \in \mathrm{gph}\, \partial f$. The *critical cone* of f at $(\bar{x}, \bar{v})$ is defined by

$$K_f(\bar{x}, \bar{v}) := \big\{ w \in \mathbb{X} \,|\, \mathrm{d}f(\bar{x})(w) = \langle \bar{v}, w \rangle \big\}. \tag{3.13}$$

When f is the indicator function of a set, this definition boils down to the classical definition of the critical cone for sets; see [8, page 109]. It is not difficult to see that the set $K_f(\bar{x}, \bar{v})$ is a cone in $\mathbb{X}$. Taking into account Proposition 2.1(ii), we conclude that the domain of the second subderivative $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is always included in the critical cone $K_f(\bar{x}, \bar{v})$ provided that $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function. The following result provides conditions under which the domain of the second subderivative is the entire critical cone.

Proposition 3.4 (domain of second subderivatives)**.** *Assume that $f : \mathbb{X} \to \overline{\mathbb{R}}$ is finite at $\bar{x}$ with $\bar{v} \in \partial^p f(\bar{x})$ and that for every $w \in K_f(\bar{x}, \bar{v})$ we have $\mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) \neq \emptyset$. Then for all $w \in K_f(\bar{x}, \bar{v})$ we have*

$$-r\|w\|^2 \leq \mathrm{d}^2 f(\bar{x}, \bar{v})(w) \leq \inf_{z \in \mathbb{X}} \big\{ \mathrm{d}^2 f(\bar{x})(w \,|\, z) - \langle z, \bar{v} \rangle \big\} < \infty, \tag{3.14}$$

where $r \in \mathbb{R}_+$ is a constant satisfying (2.1). In particular, we have $\mathrm{dom}\, \mathrm{d}^2 f(\bar{x}, \bar{v}) = K_f(\bar{x}, \bar{v})$.

Proof. The lower estimate of $d^2 f(\bar{x}, \bar{v})$ in (3.14) results from Proposition 2.1(iii), which readily implies that $d^2 f(\bar{x}, \bar{v})(0) = 0$. This tells us that the second subderivative $d^2 f(\bar{x}, \bar{v})$ is proper. Employing now Proposition 2.1(ii) gives us the inclusion $\operatorname{dom} d^2 f(\bar{x}, \bar{v}) \subset K_f(\bar{x}, \bar{v})$. The upper estimate of $d^2 f(\bar{x}, \bar{v})(w)$ in (3.14) directly comes from (3.12). By assumptions, for any $w \in K_f(\bar{x}, \bar{v})$, there exists a z_w so that $d^2 f(\bar{x})(w \,|\, z_w) < \infty$. This guarantees that the infimum term in (3.14) is finite. Pick $w \in K_f(\bar{x}, \bar{v})$ and observe from (3.14) that $d^2 f(\bar{x}, \bar{v})(w)$ is finite. This yields the inclusion $K_f(\bar{x}, \bar{v}) \subset \operatorname{dom} d^2 f(\bar{x}, \bar{v})$, which completes the proof. $\qquad \square$

The following example, taken from [21, page 636], shows the domain of the second subderivative can be the entire set $K_f(\bar{x}, \bar{v})$ even if the assumption on the domain of the parabolic subderivative in Proposition 3.4 fails. As shown in the next section, however, this condition is automatically satisfied for composite functions appearing in (1.1).

Example 3.5 (domain of second subderivative). Define the function $f : \mathbb{X} \to \overline{\mathbb{R}}$ with $\mathbb{X} = \mathbb{R}^2$ by $f(x_1, x_2) = |x_2 - x_1^{4/3}| - x_1^2$. As argued in [21, page 636], the subderivative and subdifferential of f at $\bar{x} = (0, 0)$, respectively, are

$$df(\bar{x})(w) = |w_2| \quad \text{and} \quad \partial f(\bar{x}) = \big\{ v = (v_1, v_2) \in \mathbb{R}^2 \,|\, v_1 = 0, \ |v_2| \leq 1 \big\},$$

where $w = (w_1, w_2) \in \mathbb{R}^2$. It is not hard to see that $\bar{v} = (0, 0) \in \partial^p f(\bar{x})$. Moreover, the second subderivative of f at $\bar{x}$ for $\bar{v}$ has a representation of the form

$$d^2 f(\bar{x}, \bar{v})(w) = \begin{cases} -2w_1^2 & \text{if } w_2 = 0, \\ \infty & \text{if } w_2 \neq 0. \end{cases}$$

Using the above calculation tells us that $K_f(\bar{x}, \bar{v}) = \{ w = (w_1, w_2) \,|\, w_2 = 0 \}$. Thus we have $\operatorname{dom} d^2 f(\bar{x}, \bar{v}) = K_f(\bar{x}, \bar{v})$. However, for any $w = (w_1, w_2) \in K_f(\bar{x}, \bar{v})$ with $w_1 \neq 0$ we have

$$d^2 f(\bar{x})(w \,|\, z) = \infty \quad \text{for all } z \in \mathbb{R}^2,$$

which confirms that the assumption related to the domain of the parabolic subderivative in Proposition 3.4 fails.

We proceed next by providing an important characterization of the parabolic regularity that plays a key role in our developments in this paper.

Proposition 3.6 (characterization of parabolic regularity). *Assume that $f : \mathbb{X} \to \overline{\mathbb{R}}$ is finite at $\bar{x}$ with $\bar{v} \in \partial^p f(\bar{x})$. Then the function f is parabolically regular at $\bar{x}$ for $\bar{v}$ if and only if we have*

$$d^2 f(\bar{x}, \bar{v})(w) = \inf_{z \in \mathbb{X}} \big\{ d^2 f(\bar{x})(w \,|\, z) - \langle z, \bar{v} \rangle \big\} \tag{3.15}$$

for all $w \in K_f(\bar{x}, \bar{v})$. Furthermore, for any $w \in \operatorname{dom} d^2 f(\bar{x}, \bar{v})$, there exists a $\bar{z} \in \operatorname{dom} d^2 f(\bar{x})(w \,|\, \cdot)$ such that

$$d^2 f(\bar{x}, \bar{v})(w) = d^2 f(\bar{x})(w \,|\, \bar{z}) - \langle \bar{z}, \bar{v} \rangle. \tag{3.16}$$

Proof. It follows from $\bar{v} \in \partial^p f(\bar{x})$ and Proposition 2.1(ii)-(iii) that the second subderivative $d^2 f(\bar{x}, \bar{v})$ is a proper function and

$$\operatorname{dom} d^2 f(\bar{x}, \bar{v}) \subset K_f(\bar{x}, \bar{v}). \tag{3.17}$$

Assume now that f is parabolically regular at $\bar{x}$ for $\bar{v}$. If there exists a $w \in K_f(\bar{x}, \bar{v}) \setminus \operatorname{dom} d^2 f(\bar{x}, \bar{v})$, then (3.15) clearly holds due to (3.12). Suppose now that $w \in \operatorname{dom} d^2 f(\bar{x}, \bar{v})$. By Definition 3.1, there are sequences $t_k \downarrow 0$ and $w_k \to w$ for which we have

$$\Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) \to d^2 f(\bar{x}, \bar{v})(w) \quad \text{and} \quad \limsup_{k \to \infty} \frac{\|w_k - w\|}{t_k} < \infty.$$

9

Since the sequence $z_k := 2[w_k - w]/t_k$ is bounded, we can assume by passing to a subsequence if necessary that $z_k \to \bar{z}$ as $k \to \infty$ for some $\bar{z} \in \mathbb{X}$. Thus we have $w_k = w + \frac{1}{2}t_k z_k$ and

$$
\begin{aligned}
\mathrm{d}^2 f(\bar{x}, \bar{v})(w) &= \lim_{k \to \infty} \frac{f(\bar{x} + t_k w_k) - f(\bar{x}) - t_k \langle \bar{v}, w_k \rangle}{\frac{1}{2}t_k^2} \\
&= \lim_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \langle \bar{v}, w \rangle}{\frac{1}{2}t_k^2} - \langle \bar{v}, z_k \rangle \\
&\geq \liminf_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2} - \langle \bar{v}, \bar{z} \rangle \\
&\geq \mathrm{d}^2 f(\bar{x})(w \,|\, \bar{z}) - \langle \bar{v}, \bar{z} \rangle.
\end{aligned}
$$

Combining this and (3.12) implies that (3.15) and (3.16) hold for all $w \in \operatorname{dom} \mathrm{d}^2 f(\bar{x}, \bar{v})$. To obtain the opposite implication, assume that (3.15) holds for all $w \in K_f(\bar{x}, \bar{v})$. To prove the parabolic regularity of f at $\bar{x}$ for $\bar{v}$, let $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) < \infty$, which by (3.17) yields $w \in K_f(\bar{x}, \bar{v})$. Employing now [21, Proposition 13.64] results in

$$
\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \inf_{z \in \mathbb{X}} \left\{ \mathrm{d}^2 f(\bar{x})(w \,|\, z) - \langle z, \bar{v} \rangle \right\} = \liminf_{\substack{t \downarrow 0, \, w' \to w \\ [w' - w]/t \text{ bounded}}} \Delta_t^2 f(\bar{x}, \bar{v})(w').
$$

The last equality clearly justifies (3.5), and thus f is parabolically regular at $\bar{x}$ for $\bar{v}$. This completes the proof. $\qquad\square$

We next show that the indicator function of the cone of $n \times n$ positive semidefinite symmetric matrices, denoted by $\mathbb{S}_+^n$, is parabolic regular. This can be achieved via [14, Theorem 6.2] using the theory of $\mathcal{C}^2$-cone reducible sets but below we give an independent proof via Proposition 3.6.

Example 3.7 (parabolic regularity of $\mathbb{S}_+^n$)**.** Let $\mathbb{S}_-^n$ stand for the cone of $n \times n$ negative semidefinite symmetric matrices. For any $A \in \mathbb{S}_-^n$, we are going to show that $f := \delta_{\mathbb{S}_-^n}$ is parabolic regular at A for any $V \in N_{\mathbb{S}_-^n}(A)$. Since we have $\mathbb{S}_+^n = -\mathbb{S}_-^n$, this clearly yields the same property for $\mathbb{S}_+^n$. Using the notation in Example 3.3, we can equivalently write

$$
\mathbb{S}_-^n = \left\{ A \in \mathbb{S}^n \,|\, \lambda_1(A) \leq 0 \right\}, \tag{3.18}
$$

which in turn implies that $\delta_{\mathbb{S}_-^n}(A) = \delta_{\mathbb{R}_-}(\lambda_1(A))$ for any $A \in \mathbb{S}^n$. If A is negative definite, i.e., $\lambda_1(A) < 0$, then our claim immediately follows from $N_{\mathbb{S}_-^n}(A) = \{0\}$ for this case. Otherwise, we have $\lambda_1(A) = 0$. Pick $V \in N_{\mathbb{S}_-^n}(A)$ and conclude from (3.18) and the chain rule from convex analysis that $N_{\mathbb{S}_-^n}(A) = \mathbb{R}_+ \partial \lambda_1(A)$, which implies that $V = rB$ for some $r \in \mathbb{R}_+$ and $B \in \partial \lambda_1(A)$. If $r = 0$, we get $V = 0$ and parabolic regularity of $\delta_{\mathbb{S}_-^n}$ at A for V follows directly from the definition. Assume now $r > 0$ and pick $W \in K_f(A, V)$. The latter amounts to

$$
\langle V, W \rangle = \mathrm{d}\delta_{\mathbb{S}_-^n}(A)(W) = 0 \quad \text{and} \quad W \in T_{\mathbb{S}_-^n}(A).
$$

Employing now [2, Proposition 2.61] tells us that $\mathrm{d}\lambda_1(A)(W) \leq 0$. Since $B \in \partial \lambda_1(A)$ and $\langle B, W \rangle = 0$, we arrive at $\mathrm{d}\lambda_1(A)(W) = 0$. We know from [21, Example 10.28] that

$$
\mathrm{d}\lambda_1(A)(W) = \lim_{\substack{t \downarrow 0 \\ W' \to W}} \Delta_t \lambda_1(A)(W') \quad \text{with} \quad \Delta_t \lambda_1(A)(W') := \frac{\lambda_1(A + tW') - \lambda_1(A)}{t}.
$$

Using direct calculations, we conclude for any $t > 0$ and $W' \in \mathbb{S}^n$ that

$$
\Delta_t^2 \delta_{\mathbb{S}_-^n}(A, V)(W') = \Delta_t^2 \delta_{\mathbb{R}_-}(\lambda_1(A), r)(\Delta_t \lambda_1(A)(W')) + r \Delta_t^2 \lambda_1(A, B)(W'),
$$

10

which in turn results in

$$\mathrm{d}^2\delta_{\mathbb{S}^n_-}(A,V)(W) \geq \mathrm{d}^2\delta_{\mathbb{R}_-}(\lambda_1(A),r)(\mathrm{d}\lambda_1(A)(W)) + r\mathrm{d}^2\lambda_1(A,B)(W).$$

Since $r > 0$, $\lambda_1(A) = 0$, and $\mathrm{d}\lambda_1(A)(W) = 0$, we conclude from [14, Example 3.4] that

$$\mathrm{d}^2\delta_{\mathbb{R}_-}(\lambda_1(A),r)(\mathrm{d}\lambda_1(A)(W)) = \delta_{K_{\mathbb{R}_-}(\lambda_1(A),r)}(0) = \delta_{\{0\}}(0) = 0.$$

Using this together with (3.8) brings us to

$$\mathrm{d}^2\delta_{\mathbb{S}^n_-}(A,V)(W) \geq -2r\langle B, WA^\dagger W\rangle = -2\langle V, WA^\dagger W\rangle.$$

On the other hand, we conclude from (3.14) that

$$\mathrm{d}^2\delta_{\mathbb{S}^n_-}(A,V)(W) \leq -\sigma_{T^2_{\mathbb{S}^n_-}(A,W)}(V) = -2\langle V, WA^\dagger W\rangle,$$

where the last equality comes from [2, page 487] with $\sigma_{T^2_{\mathbb{S}^n_-}(A,W)}$ standing for the support function of $T^2_{\mathbb{S}^n_-}(A,W)$. Combining these confirms that

$$\mathrm{d}^2\delta_{\mathbb{S}^n_-}(A,V)(W) = -\sigma_{T^2_{\mathbb{S}^n_-}(A,W)}(V) = -2\langle V, WA^\dagger W\rangle \quad \text{for all } W \in K_f(A,V).$$

This together with Proposition 3.6 tells us that $\mathbb{S}^n_-$ is parabolic regular at A for V.

We are now in a position to establish the main result of this section, which states that parabolically regular functions are always twice epi-differentiable.

Theorem 3.8 (twice epi-differenitability of parabolically regular functions). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ be finite at $\bar{x}$ and $\bar{v} \in \partial^p f(\bar{x})$ and let f be parabolically epi-differentiable at $\bar{x}$ for every $w \in K_f(\bar{x},\bar{v})$. If f is parabolically regular at $\bar{x}$ for $\bar{v}$, then it is properly twice epi-differentiable at $\bar{x}$ for $\bar{v}$ with*

$$\mathrm{d}^2 f(\bar{x},\bar{v})(w) = \begin{cases} \min_{z\in\mathbb{X}} \left\{ \mathrm{d}^2 f(\bar{x})(w\,|\,z) - \langle z, \bar{v}\rangle \right\} & \text{if } w \in K_f(\bar{x},\bar{v}), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.19}$$

Proof. It follows from the parabolic epi-differentiability of f at $\bar{x}$ for every $w \in K_f(\bar{x},\bar{v})$ and Proposition 3.4 that $\mathrm{dom}\,\mathrm{d}^2 f(\bar{x},\bar{v}) = K_f(\bar{x},\bar{v})$. This together with (3.15) and (3.16) justifies the second subderivative formula (3.19). To establish the twice epi-differentiability of f at $\bar{x}$ for $\bar{v}$, we are going to show that (2.3) holds for all $w \in \mathbb{X}$. Pick $w \in K_f(\bar{x},\bar{v})$ and an arbitrary sequence $t_k \downarrow 0$. Since f is parabolically regular at $\bar{x}$ for $\bar{v}$, by Proposition 3.6, we find a $\bar{z} \in \mathbb{X}$ such that

$$\mathrm{d}^2 f(\bar{x},\bar{v})(w) = \mathrm{d}^2 f(\bar{x})(w\,|\,\bar{z}) - \langle \bar{z}, \bar{v}\rangle. \tag{3.20}$$

By the parabolic epi-differentiability of f at $\bar{x}$ for w, we find a sequence $z_k \to \bar{z}$ for which we have

$$\mathrm{d}^2 f(\bar{x})(w\,|\,\bar{z}) = \lim_{k\to\infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k\mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2}.$$

Define $w_k := w + \frac{1}{2}t_k z_k$ for all $k \in \mathbb{N}$. Using this and $w \in K_f(\bar{x},\bar{v})$, we obtain

$$\begin{aligned} \Delta^2_{t_k} f(\bar{x},\bar{v})(w_k) &= \frac{f(\bar{x}+t_k w_k) - f(\bar{x}) - t_k\langle \bar{v}, w_k\rangle}{\frac{1}{2}t_k^2} \\ &= \frac{f(\bar{x}+t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k\mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2} - \langle \bar{v}, z_k\rangle. \end{aligned}$$

11

This together with (3.20) results in

$$\lim_{k \to \infty} \Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) = \mathrm{d}^2 f(\bar{x})(w \mid \bar{z}) - \langle \bar{v}, \bar{z} \rangle = \mathrm{d}^2 f(\bar{x}, \bar{v})(w),$$

which justifies (2.3) for every $w \in K_f(\bar{x}, \bar{v})$. Finally, we are going to show the validity of (2.3) for every $w \notin K_f(\bar{x}, \bar{v})$. For any such a w, we conclude from (3.19) that $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \infty$. Pick an arbitrary sequence $t_k \downarrow 0$ and set $w_k := w$ for all $k \in \mathbb{N}$. Thus we have

$$\infty = \mathrm{d}^2 f(\bar{x}, \bar{v})(w) \le \liminf_{k \to \infty} \Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) \le \limsup_{t \downarrow 0} \Delta_{t_k}^2 f(\bar{x}, \bar{v})(w_k) \le \infty = \mathrm{d}^2 f(\bar{x}, \bar{v})(w),$$

which again proves (2.3) for all $w \notin K_f(\bar{x}, \bar{v})$. This completes the proof of the Theorem. $\qquad \square$

The above theorem provides a very important generalization of a similar result obtained recently by the authors and Mordukhovich in [14, Theorem 3.6] in which the twice epi-differentiability of set indicator functions was established. It is not hard to see that the assumptions of Theorem 3.8 boils down to those in [14, Theorem 3.6]. To the best of our knowledge, the only results related to the twice epi-differentiability of functions, beyond set indicator functions, are [21, Theorem 13.14] and [24, Theorem 3.1] in which this property was proven for the fully amenable and eigenvalue functions, respectively. We will derive these results in Section 5 as an immediate consequence of our chain rule for the second subderivative.

We proceed with an important consequence of Theorem 3.8 in which the proto-differentiability of subgradient mappings is established under parabolic regularity. Recall that a set-valued mapping $S : \mathbb{X} \rightrightarrows \mathbb{Y}$ is said to be *proto-differentiable* at $\bar{x}$ for $\bar{y}$ with $(\bar{x}, \bar{y}) \in \mathrm{gph}\, S$ if the set $\mathrm{gph}\, S$ is geometrically derivable at $(\bar{x}, \bar{y})$. When this condition holds for the set-valued mapping S at $\bar{x}$ for $\bar{y}$, we refer to $DS(\bar{x}, \bar{y})$ as the *proto-derivative* of S at $\bar{x}$ for $\bar{y}$. The connection between the twice epi-differentiablity of a function and the proto-differentiability of its subgradient mapping was observed first by Rockafellar in [20] for convex functions and was extended later in [16] for prox-regular functions. Recall that a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is called prox-regular at $\bar{x}$ for $\bar{v}$ if f is finite at $\bar{x}$ and is locally l.s.c. around $\bar{x}$ with $\bar{v} \in \partial f(\bar{x})$ and there are constant $\varepsilon > 0$ and $r \ge 0$ such that for all $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $f(x) \le f(\bar{x}) + \varepsilon$ we have

$$f(x) \ge f(u) + \langle v, x - u \rangle - \tfrac{r}{2} \|x - u\|^2 \qquad \text{for all} \quad (u, v) \in (\mathrm{gph}\, \partial f) \cap \mathbb{B}_\varepsilon(\bar{x}, \bar{v}).$$

Moreover, we say that f is subdifferentially continuous at $\bar{x}$ for $\bar{v}$ if $(x_k, v_k) \to (\bar{x}, \bar{v})$ with $v_k \in \partial f(x_k)$, one has $f(x_k) \to f(\bar{x})$.

Corollary 3.9 (proto-differentiability under parabolic regularity)**.** *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ be prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{v}$ and let f be parabolically epi-differentiable at $\bar{x}$ for every $w \in K_f(\bar{x}, \bar{v})$. If f is parabolically regular at $\bar{x}$ for $\bar{v}$, then the following equivalent conditions hold:*

   (i) *the function f is twice epi-differentiable at $\bar{x}$ for $\bar{v}$;*
  (ii) *the subgradient mapping ∂f is proto-differentiable at $\bar{x}$ for $\bar{v}$.*
Furthermore, the proto-derivative of the subgradient mapping ∂f at $\bar{x}$ for $\bar{v}$ can be calculated by

$$D(\partial f)(\bar{x}, \bar{v})(w) = \partial \big(\tfrac{1}{2} \mathrm{d}^2 f(\bar{x}, \bar{v}) \big)(w) \quad \text{for all} \ w \in \mathbb{X}. \tag{3.21}$$

Proof. Note that $\bar{v} \in \partial^p f(\bar{x})$ since f is prox-regular at $\bar{x}$ for $\bar{v}$. Employing now Theorem 3.8 gives us (i). The equivalence between (i) and (ii) and the validity of (3.21) come from [21, Theorem 13.40]. $\qquad \square$

4 Variational Properties of Parabolic Subderivatives

This section is devoted to second-order analysis of parabolic subderivatives of extended-real-valued functions that are locally Lipschitz continuous relative to their domains. We pay special attention to functions that are expressed as a composition of a convex function and a twice differentiable function. We begin with the following result that gives us sufficient conditions for finding the domain of the parabolic subderivative.

Proposition 4.1 (properties of parabolic subderivatives). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ be finite at $\bar{x}$ and let f be Lipschitz continuous around $\bar{x}$ relative to its domain with constant $\ell \in \mathbb{R}_+$. Assume that $w \in T_{\operatorname{dom} f}(\bar{x})$ and that f is parabolic epi-differentiable at $\bar{x}$ for w. Then the following conditions hold:*

   (i) $\operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) = T^2_{\operatorname{dom} f}(\bar{x}, w)$;
   (ii) $\operatorname{dom} f$ *is parabolically derivable at $\bar{x}$ for w.*

Proof. Since $w \in T_{\operatorname{dom} f}(\bar{x})$, we conclude from Proposition 2.2 that $\mathrm{d}f(\bar{x})(w)$ is finite. To prove (i), observe first that by definition, we always have the inclusion

$$\operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) \subset T^2_{\operatorname{dom} f}(\bar{x}, w). \tag{4.1}$$

To obtain the opposite inclusion, take $z \in T^2_{\operatorname{dom} f}(\bar{x}, w)$. This tells us that there exist sequences $t_k \downarrow 0$ and $z_k \to z$ so that $\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k \in \operatorname{dom} f$. Since f is parabolically epi-differentiable at $\bar{x}$ for w, we have $\operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) \neq \emptyset$. Thus there exists a $z_w \in \mathbb{X}$ such that $\mathrm{d}^2 f(\bar{x})(w \,|\, z_w) < \infty$. Moreover, corresponding to the sequence t_k, we find another sequence $z_k' \to z_w$ such that

$$\mathrm{d}^2 f(\bar{x})(w \,|\, z_w) = \lim_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k') - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2} t_k^2}.$$

Since $\mathrm{d}^2 f(\bar{x})(w \,|\, z_w) < \infty$, we can assume without loss of generality that $\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k' \in \operatorname{dom} f$ for all $k \in \mathbb{N}$. Using these together with the Lipschitz continuity of f around $\bar{x}$ relative to its domain, we have for all k sufficiently large that

$$
\begin{aligned}
\frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2} t_k^2}
&= \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k') - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2} t_k^2} \\
&\quad + \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k) - f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k')}{\frac{1}{2} t_k^2} \\
&\leq \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k') - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2} t_k^2} \\
&\quad + \ell \| z_k - z_k' \|.
\end{aligned}
$$

Passing to the limit results in the inequality

$$\mathrm{d}^2 f(\bar{x})(w \,|\, z) \leq \mathrm{d}^2 f(\bar{x})(w \,|\, z_w) + \ell \| z - z_w \|, \tag{4.2}$$

which in turn yields $\mathrm{d}^2 f(\bar{x})(w \,|\, z) < \infty$, i.e., $z \in \operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot)$. This justifies the opposite inclusion in (4.1) and hence proves (i).

Turning now to (ii), we conclude from (4.1) and the parabolic epi-differentiability of f at $\bar{x}$ for w that the second-order tangent set $T^2_{\operatorname{dom} f}(\bar{x}, w)$ is nonempty. Moreover, it follows from [21, Example 13.62(b)] that the parabolic epi-differentiability of f at $\bar{x}$ for w yields the parabolic derivability of $\operatorname{epi} f$ at $(\bar{x}, f(\bar{x}))$ for $(w, \mathrm{d}f(\bar{x})(w))$. The latter clearly enforces the same property for $\operatorname{dom} f$ at $\bar{x}$ for w and hence completes the proof. □

13

It is important to notice the parabolic epi-differentiability of f in Proposition 4.1 is essential to ensure that condition (i) therein, namely the characterization of the domain of the parabolic subderivative, is satisfied. Indeed, as mentioned in the proof of this proposition, inclusion (4.1) always holds. If the latter condition fails, this inclusion can be strict. For example, the function f from Example 3.5 is not parabolic epi-differentiable at $\bar{x} = (0,0)$ for any vector $w = (w_1, w_2) \in K_f(\bar{x}, \bar{v})$ with $w_1 \neq 0$ since $\operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) = \emptyset$. On the other hand, we have $\operatorname{dom} f = \mathbb{R}^2$ and thus $T^2_{\operatorname{dom} f}(\bar{x}, w) = \mathbb{R}^2$ for any such a vector $w \in K_f(\bar{x}, \bar{v})$, and so condition (i) in Proposition 4.1 fails.

Given a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ finite at $\bar{x}$, in the rest of this paper, we mainly focus on the case when this function has a representation of the form

$$f(x) = (g \circ F)(x) \quad \text{for all } x \in \mathcal{O}, \tag{4.3}$$

where $\mathcal{O}$ is a neighborhood of $\bar{x}$ and where the functions F and g are satisfying the following conditions:

- $F : \mathbb{X} \to \mathbb{Y}$ is twice differentiable at $\bar{x}$;
- $g : \mathbb{Y} \to \overline{\mathbb{R}}$ is proper, l.s.c., convex, and Lipschitz continuous around $F(\bar{x})$ relative to its domain with constant $\ell \in \mathbb{R}_+$.

It is not hard to see that the imposed assumptions on g from representation (4.3) implies that $\operatorname{dom} g$ is locally closed around $F(\bar{x})$, namely for some $\varepsilon > 0$ the set $(\operatorname{dom} g) \cap \mathbb{B}_\varepsilon(F(\bar{x}))$ is closed. Taking the neighborhood $\mathcal{O}$ from (4.3), we obtain

$$(\operatorname{dom} f) \cap \mathcal{O} = \{ x \in \mathcal{O} \,|\, F(x) \in \operatorname{dom} g \}. \tag{4.4}$$

It has been well understood that the second-order variational analysis of the composite form (4.3) requires a certain qualification condition. The following definition provides the one we utilize in this paper.

Definition 4.2 (metric subregularity constraint qualification). *Assume that the function $f : \mathbb{X} \to \overline{\mathbb{R}}$ has representation* (4.3) *around* $\bar{x} \in \operatorname{dom} f$. *We say that the metric subregularity constraint qualification holds for the constraint set* (4.4) *at* $\bar{x}$ *if there exist a constants* $\kappa \in \mathbb{R}_+$ *and a neighborhood* U *of* $\bar{x}$ *such that*

$$d(x, \operatorname{dom} f) \leq \kappa \, d(F(x), \operatorname{dom} g) \quad \text{for all } x \in U. \tag{4.5}$$

By definition, the metric subregularity constraint qualification for the constraint set (4.4) at $\bar{x}$ amounts to the metric subregularity of the mapping $x \mapsto F(x) - \operatorname{dom} g$ at $(\bar{x}, 0)$. The more traditional and well known constraint qualification for (4.3) is

$$\partial^\infty g(F(\bar{x})) \cap \ker \nabla F(\bar{x})^* = \{0\},$$

where $\partial^\infty \varphi(\bar{x})$ stands for the *singular subdifferential* of $\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom} \varphi$ defined by

$$\partial^\infty \varphi(\bar{x}) := \{ v \in \mathbb{R}^n \,|\, (v, 0) \in N_{\operatorname{epi} \varphi}(\bar{x}, \varphi(\bar{x})) \}.$$

By the coderivative criterion from [12, Theorem 3.3], the latter constraint qualification amounts to the metric regularity of the mapping $(x, \alpha) \mapsto (F(x), \alpha) - \operatorname{epi} g$ around $((\bar{x}, g(F(\bar{x}))), (0, 0))$. Since g is convex, by [12, Proposition 1.25], the aforementioned constraint qualification can be equivalently written as

$$N_{\operatorname{dom} g}(F(\bar{x})) \cap \ker \nabla F(\bar{x})^* = \{0\}, \tag{4.6}$$

which by the coderivative criterion from [12, Theorem 3.3] is equivalent to the metric regularity of the mapping $x \mapsto F(x) - \operatorname{dom} g$ around $(\bar{x}, 0)$. Thus, for the composite form (4.3), the

14

metric regularity of $(x, \alpha) \mapsto (F(x), \alpha) - \operatorname{epi} g$ around $\big((\bar{x}, g(F(\bar{x}))), (0,0)\big)$ amounts to that of $x \mapsto F(x) - \operatorname{dom} g$ around $(\bar{x}, 0)$. This may not be true if the metric regularity is replaced with the metric subregularity; see [13, Proposition 3.1] and [13, Example 3.3] for more details and discussions about these conditions. So, it is worth reiterating that instead of using the metric subregularity of the *epigraphical* mapping, our second-order variational analysis will be carried out under a weaker (and simpler) metric subregularity of the *domain* mapping.

As observed recently in [13], (4.5) suffices to conduct first- and second-order variational analysis of (4.3) when the convex function g therein is merely piecewise linear-quadratic. In what follows, we will show using a different approach that such results can be achieved for (4.3) as well. We continue our analysis by recalling the following first- and second-order chain rules, obtained recently in [13, 14].

Proposition 4.3 (first- and second-order chain rules). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ have the composite representation* (4.3) *at $\bar{x} \in \operatorname{dom} f$ and $\bar{v} \in \partial f(\bar{x})$ and let the metric subregularity constraint qualification hold for the constraint set* (4.4) *at $\bar{x}$. Then the following hold:*

(i) *for any $w \in \mathbb{X}$, the following subderivative chain rule for f at $\bar{x}$ holds:*

$$\mathrm{d}f(\bar{x})(w) = \mathrm{d}g(F(\bar{x}))(\nabla F(\bar{x})w);$$

(ii) *we have the chain rules*

$$\partial^p f(\bar{x}) = \partial f(\bar{x}) = \nabla F(\bar{x})^* \partial g(F(\bar{x})) \quad and \quad T_{\operatorname{dom} f}(\bar{x}) = \big\{ w \in \mathbb{X} \mid \nabla F(\bar{x})w \in T_{\operatorname{dom} g}(F(\bar{x})) \big\}.$$

If, in addition, $w \in T_{\operatorname{dom} f}(\bar{x})$ and the function g from (4.3) *is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x})w$, then we have*

$$z \in T^2_{\operatorname{dom} f}(\bar{x}, w) \iff \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w, w) \in T^2_{\operatorname{dom} g}\big(F(\bar{x}), \nabla F(\bar{x})w\big). \tag{4.7}$$

Moreover, $\operatorname{dom} f$ is parabolically derivable at $\bar{x}$ for w.

Proof. The subderivative chain rule in (i) was established recently in [13, Theorem 3.4]. The subdifferential chain rule in (ii) was taken from [13, Theorem 3.6]. As mentioned in Section 2, the inclusion $\partial^p f(\bar{x}) \subset \partial f(\bar{x})$ always holds. The opposite inclusion can be justified using the aforementioned subdifferential chain rule and the convexity of g; see [13, Theorem 4.4] for a similar result. The chain rule for the tangent cone to $\operatorname{dom} f$ at $\bar{x}$ results from Proposition 2.2 and the subderivative chain rule for f at $\bar{x}$ in (i). If in addition $w \in T_{\operatorname{dom} f}(\bar{x})$ and g is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x})w$, then it follows from Proposition 4.1 that $\operatorname{dom} g$ is parabolically derivable at $F(\bar{x})$ for $\nabla F(\bar{x})w$. Appealing now to [14, Theorem 4.5] implies that $\operatorname{dom} f$ is parabolically derivable at $\bar{x}$ for w. Finally, the chain rule (4.7) was taken from [14, Theorem 4.5]. This completes the proof. $\square$

We continue by establishing a chain rule for the parabolic subderivative, which is important for our developments in the next section.

Theorem 4.4 (chain rule for parabolic subderivatives). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ have the composite representation* (4.3) *at $\bar{x} \in \operatorname{dom} f$ and $w \in T_{\operatorname{dom} f}(\bar{x})$ and let the metric subregularity constraint qualification hold for the constraint set* (4.4) *at $\bar{x}$. Assume that the function g from* (4.3) *is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x})w$. Then the following conditions are satisfied:*

(i) *for any $z \in \mathbb{X}$ we have*

$$\mathrm{d}^2 f(\bar{x})(w \mid z) = \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \mid \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w, w)); \tag{4.8}$$

(ii) *the domain of the parabolic subderivative of f at $\bar{x}$ for w is given by*

$$\operatorname{dom} \operatorname{d}^2 f(\bar{x})(w \,|\, \cdot) = T^2_{\operatorname{dom} f}(\bar{x}, w);$$

(iii) *f is parabolically epi-differentiable at $\bar{x}$ for w.*

Proof. Pick $z \in \mathbb{X}$ and set $u := \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w, w)$. We prove (i)-(iii) in a parallel way. Assume that $z \notin T^2_{\operatorname{dom} f}(\bar{x}, w)$. As mentioned in the proof of Proposition 4.1, inclusion (4.1) always holds. This implies that $\operatorname{d}^2 f(\bar{x})(w \,|\, z) = \infty$. On the other hand, by (4.7) we get $u \notin T^2_{\operatorname{dom} g}\big(F(\bar{x}), \nabla F(\bar{x})w\big)$. Employing Proposition 4.1(i) for the function g and $\nabla F(\bar{x})w \in T_{\operatorname{dom} g}(F(\bar{x}))$ gives us

$$\operatorname{dom} \operatorname{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, \cdot) = T^2_{\operatorname{dom} g}(F(\bar{x}), \nabla F(\bar{x})w). \tag{4.9}$$

Combining these tells us that $\operatorname{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) = \infty$, which in turn justifies (4.8) for every $z \notin T^2_{\operatorname{dom} f}(\bar{x}, w)$. Consider an arbitrary sequence $t_k \downarrow 0$ and set $z_k := z$ for all $k \in \mathbb{N}$. Then we have

$$
\begin{aligned}
\operatorname{d}^2 f(\bar{x})(w \,|\, z) &\leq \liminf_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k) - f(\bar{x}) - t_k \operatorname{d} f(\bar{x})(w)}{\frac{1}{2} t_k^2} \\
&\leq \limsup_{k \to \infty} \frac{f(\bar{x} + t_k w + \frac{1}{2} t_k^2 z_k) - f(\bar{x}) - t_k \operatorname{d} f(\bar{x})(w)}{\frac{1}{2} t_k^2} \\
&\leq \infty = \operatorname{d}^2 f(\bar{x})(w \,|\, z),
\end{aligned}
$$

which in turn justifies (3.11) for all $z \notin T^2_{\operatorname{dom} f}(\bar{x}, w)$.

Since g is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x})w$, Proposition 4.1(ii) tells us that $\operatorname{dom} g$ is parabolically derivable at $F(\bar{x})$ for $\nabla F(\bar{x})w$. We conclude from Proposition 4.3 that $\operatorname{dom} f$ is parabolically derivable at $\bar{x}$ for w. In particular, we have

$$T^2_{\operatorname{dom} f}(\bar{x}, w) \neq \emptyset. \tag{4.10}$$

Pick now $z \in T^2_{\operatorname{dom} f}(\bar{x}, w)$ and then consider an arbitrary sequence $t_k \downarrow 0$. Thus, by definition, for the aforementioned sequence t_k, we find a sequence $z_k \to z$ as $k \to \infty$ such that

$$x_k := \bar{x} + t_k w + \frac{1}{2} t_k^2 z_k \in \operatorname{dom} f \quad \text{for all} \;\; k \in \mathbb{N}. \tag{4.11}$$

Moreover, since g is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x})w$, we find a sequence $u_k \to u$ such that

$$\operatorname{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) = \lim_{k \to \infty} \frac{g(F(\bar{x}) + t_k \nabla F(\bar{x})w + \frac{1}{2} t_k^2 u_k) - g(F(\bar{x})) - t_k \operatorname{d} g(F(\bar{x}))(\nabla F(\bar{x})w)}{\frac{1}{2} t_k^2}.$$

$$\tag{4.12}$$

It follows from (4.7) that $u \in T^2_{\operatorname{dom} g}\big(F(\bar{x}), \nabla F(\bar{x})w\big)$. Combining this with (4.9) tells us that $\operatorname{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) < \infty$. This implies that $y_k := F(\bar{x}) + t_k \nabla F(\bar{x})w + \frac{1}{2} t_k^2 u_k \in \operatorname{dom} g$ for all k sufficiently large. Remember that g is Lipschitz continuous around $F(\bar{x})$ relative to its domain

with constant ℓ. Using this together with Proposition 4.3(i), (4.11), and (4.12), we obtain

$$
\begin{aligned}
\mathrm{d}^2 f(\bar{x})(w \,|\, z) \;\leq\; & \liminf_{k\to\infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2} \\
\leq\; & \limsup_{k\to\infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2} \\
=\; & \limsup_{k\to\infty} \frac{g(F(x_k)) - g(F(\bar{x})) - t_k \mathrm{d}g(F(\bar{x}))(\nabla F(\bar{x})w)}{\frac{1}{2}t_k^2} \\
\leq\; & \lim_{k\to\infty} \frac{g(y_k) - g(F(\bar{x})) - t_k \mathrm{d}g(F(\bar{x}))(\nabla F(\bar{x})w)}{\frac{1}{2}t_k^2} + \limsup_{k\to\infty} \frac{g(F(x_k)) - g(y_k)}{\frac{1}{2}t_k^2} \\
\leq\; & \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) + \limsup_{k\to\infty} \ell \| \nabla F(\bar{x})z_k + \nabla^2 F(\bar{x})(w,w) - u_k + \frac{o(t_k^2)}{t_k^2} \| \\
=\; & \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u).
\end{aligned} \tag{4.13}
$$

On the other hand, it is not hard to see that for any $z \in \mathbb{X}$, we always have

$$
\mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) \leq \mathrm{d}^2 f(\bar{x})(w \,|\, z).
$$

Combining this and (4.13) implies that

$$
\mathrm{d}^2 f(\bar{x})(w \,|\, z) = \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u)
$$

and that

$$
\mathrm{d}^2 f(\bar{x})(w \,|\, z) = \lim_{k\to\infty} \frac{f(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2}.
$$

These prove (4.8) and (3.11) for any $z \in T^2_{\mathrm{dom}\, f}(\bar{x}, w)$, respectively, and hence we finish the proof of (i).

Next, we are going to verify (ii). We already know that inclusion (4.1) always holds. To derive the opposite inclusion, pick $z \in T^2_{\mathrm{dom}\, f}(\bar{x}, w)$, which amounts to $u \in T^2_{\mathrm{dom}\, g}(F(\bar{x}), \nabla F(\bar{x})w)$ due to (4.7). By (i) and (4.9), we obtain

$$
\mathrm{d}^2 f(\bar{x})(w \,|\, z) = \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u) < \infty.
$$

This tells us that $z \in \mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot)$ and hence completes the proof of (ii).

Finally, to justify (iii), we require to prove the fulfillment of (3.11) for all $z \in \mathbb{X}$ and to show that $\mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) \neq \emptyset$. The former was proven above and so we proceed with the proof of the latter. This, indeed, follows from (4.10) and the characterization of $\mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot)$, achieved in (ii), and thus completes the proof. $\qquad\square$

It is worth mentioning that a chain rule for parabolic subderivatives for the composite form (4.3) was achieved in [21, Exercise 13.63] and [2, Proposition 3.42] when g is merely a proper l.s.c. function and the basic constraint qualification (4.6) is satisfied. Replacing the latter condition with the significantly weaker condition (4.5), we can achieve a similar result if we assume further that g is convex and locally Lipschitz continuous relative to its domain. Another important difference between Theorem 4.4 and those mentioned above is that the chain rule (4.8) obtained in [2,21] does not require the parabolic epi-differentiability of g. Indeed, the usage of the basic constraint qualification (4.6) in [2,21] allows to achieve (4.8) via a chain rule for the epigraphs of f and g similar to the one in (4.7), which is not conceivable under (4.5). These extra assumptions

on g automatically fulfill in many important composite and constrained optimization problems and so do not seem to be restrictive in our developments.

We continue by establishing two important properties for parabolic subderivatives that play crucial roles in our developments in the next section. One notable difference between the following results and those obtained in Proposition 4.1 and Theorem 4.4 is that we require the parabolic subderivative be proper. This can be achieved if the parabolic subderivative is bounded below. In general, we may not be able to guarantee this. It turns out, however, that if the vector w in the pervious results is taken from the critical cone to the function in question, which is a subset of the tangent cone to the domain of that function, this can be accomplished via (3.14). Since we only conduct our analysis in the next section over the critical cone, this will provide no harm. Below, we first show that the parabolic subderivative of an extended-real-valued function, which is locally Lipschitz continuous relative to its domain, is Lipschitz continuous relative to its domain.

Proposition 4.5 (Lipschitz continuity of of parabolic subderivatives). *Let $\psi : \mathbb{X} \to \overline{\mathbb{R}}$ be finite at $\bar{x}$ and $\bar{v} \in \partial^p \psi(\bar{x})$, and let ψ be Lipschitz continuous around $\bar{x}$ relative to its domain with constant $\ell \in \mathbb{R}_+$. Assume that $w \in K_\psi(\bar{x}, \bar{v})$ and that ψ is parabolically epi-differentiable at $\bar{x}$ for w. Then the parabolic subderivative $\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot)$ is proper, l.s.c., and Lipschitz continuous relative to its domain with constant ℓ.*

Proof. Since ψ is parabolically epi-differentiable at $\bar{x}$ for w, we get $\mathrm{dom}\,\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot) \neq \emptyset$. Let $z \in \mathrm{dom}\,\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot)$. By Proposition 3.4, we find $r \in \mathbb{R}_+$ such that

$$- r\|w\|^2 \leq \mathrm{d}^2\psi(\bar{x} \,|\, \bar{v})(w) \leq \mathrm{d}^2\psi(\bar{x})(w \,|\, z) - \langle z, \bar{v}\rangle. \tag{4.14}$$

This tells us that $\mathrm{d}^2\psi(\bar{x})(w \,|\, z)$ is finite for every $z \in \mathrm{dom}\,\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot)$ and thus the parabolic subderivative $\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot)$ is proper. Pick now $z_i \in \mathrm{dom}\,\mathrm{d}^2\psi(\bar{x})(w \,|\, \cdot)$ for $i = 1, 2$. By Proposition 4.1(i), we have $z_i \in T^2_{\mathrm{dom}\,\psi}(\bar{x}, w)$ for $i = 1, 2$. Letting $z := z_1$ and $z_w := z_2$ in (4.2) results in

$$\mathrm{d}^2\psi(\bar{x})(w \,|\, z_1) \leq \mathrm{d}^2\psi(\bar{x})(w \,|\, z_2) + \ell\|z_1 - z_2\|.$$

Similarly, we can let $z := z_2$ and $z_w := z_1$ in (4.2) and obtain

$$\mathrm{d}^2\psi(\bar{x})(w \,|\, z_2) \leq \mathrm{d}^2\psi(\bar{x})(w \,|\, z_1) + \ell\|z_1 - z_2\|.$$

Combining these implies that the parabolic subderivative is Lipschitz continuous relative to its domain. By [21, Proposition 13.64], the parabolic subderivative is always an l.s.c. function, which completes the proof. □

We end this section by obtaining an exact formula for the conjugate function of the parabolic subderivative of a convex function.

Proposition 4.6 (conjugate of parabolic subderivatives). *Let $\psi : \mathbb{X} \to \overline{\mathbb{R}}$ be an l.s.c. convex function with $\psi(\bar{x})$ finite, $\bar{v} \in \partial\psi(\bar{x})$, and $w \in K_\psi(\bar{x}, \bar{v})$. Define the function φ by $\varphi(z) := \mathrm{d}^2\psi(\bar{x})(w \,|\, z)$ for any $z \in \mathbb{X}$. If ψ is parabolically epi-differentiable at $\bar{x}$ for w and parabolically regular at $\bar{x}$ for every $v \in \partial\psi(\bar{x})$, then φ is a proper, l.s.c., and convex function and its conjugate function is given by*

$$\varphi^*(v) = \begin{cases} -\mathrm{d}^2\psi(\bar{x}, v)(w) & \text{if } v \in \mathcal{A}(\bar{x}, w), \\ \infty & \text{otherwise,} \end{cases} \tag{4.15}$$

where $\mathcal{A}(\bar{x}, w) := \{v \in \partial\psi(\bar{x}) | \, \mathrm{d}\psi(\bar{x})(w) = \langle v, w\rangle\}$.

Proof. It follows from [21, Proposition 13.64] that φ is l.s.c. Using similar arguments as the beginning of the proof of Proposition 4.5 together with (4.14) tells us that φ is proper. Also we deduce from [21, Example 13.62] that

$$\mathrm{epi}\,\varphi = T^2_{\mathrm{epi}\,\psi}\big((\bar{x}, \psi(\bar{x})), (w, \mathrm{d}\psi(\bar{x})(w))\big),$$

and thus the parabolic epi-differentiability of ψ at $\bar{x}$ for w amounts to the parabolic derivability of $\mathrm{epi}\,\psi$ at $(\bar{x}, \psi(\bar{x}))$ for $(w, \mathrm{d}\psi(\bar{x})(w))$. The latter combined with the convexity of ψ tells us that $\mathrm{epi}\,\varphi$ is a convex set in $\mathbb{X} \times \mathbb{R}$ and so φ is convex.

To verify (4.15), pick $v \in \mathcal{A}(\bar{x}, w)$. This yields $v \in \partial\psi(\bar{x}) = \partial^p\psi(\bar{x})$ and $w \in K_\psi(\bar{x}, v)$, namely the critical cone of ψ at $(\bar{x}, v)$. Using Proposition 3.6 and parabolic regularity of ψ at $\bar{x}$ for v implies that

$$\mathrm{d}^2\psi(\bar{x}, v)(w) = \inf_{z \in \mathbb{X}} \big\{ \mathrm{d}^2\psi(\bar{x})(w \,|\, z) - \langle z, v \rangle \big\} = -\varphi^*(v),$$

which clearly proves (4.15) in this case. Assume now that $v \notin \mathcal{A}(\bar{x}, w)$. This means either $v \notin \partial\psi(\bar{x})$ or $\mathrm{d}\psi(\bar{x})(w) \neq \langle v, w \rangle$. Define the parabolic difference quotients for ψ at $\bar{x}$ for w by

$$\vartheta_t(z) = \frac{\psi(\bar{x} + tw + \frac{1}{2}t^2 z) - \psi(\bar{x}) - t\mathrm{d}\psi(\bar{x})(w)}{\frac{1}{2}t^2}, \quad z \in \mathbb{X}, \;\; t > 0.$$

It is not hard to see that ϑ_t are proper, convex, and

$$\vartheta_t^*(v) = \frac{\psi(\bar{x}) + \psi^*(v) - \langle v, \bar{x} \rangle}{\frac{1}{2}t^2} + \frac{\mathrm{d}\psi(\bar{x})(w) - \langle v, w \rangle}{\frac{1}{2}t}, \quad v \in \mathbb{X}.$$

Remember that by [21, Definition 13.59] the parabolic epi-differentiability of ψ at $\bar{x}$ for w amounts to the sets $\mathrm{epi}\,\vartheta_t$ converging to $\mathrm{epi}\,\varphi$ as $t \downarrow 0$ and that the functions ϑ_t and φ are proper, l.s.c. and convex. Appealing to [21, Theorem 11.34] tells us that the former is equivalent to the sets $\mathrm{epi}\,\vartheta_t^*$ converging to $\mathrm{epi}\,\varphi^*$ as $t \downarrow 0$. This, in particular, means that for any $v \notin \mathcal{A}(\bar{x}, w)$ and any sequence $t_k \downarrow 0$, there exists a sequence $v_k \to v$ such that

$$\varphi^*(v) = \lim_{k \to \infty} \vartheta_{t_k}^*(v_k).$$

If $v \notin \partial\psi(\bar{x})$, then we have

$$\psi(\bar{x}) + \psi^*(v) - \langle v, \bar{x} \rangle > 0.$$

Since ψ^* is l.s.c., we get

$$\liminf_{k \to \infty} \frac{\psi(\bar{x}) + \psi^*(v_k) - \langle v_k, \bar{x} \rangle}{\frac{1}{2}t_k} + \frac{\mathrm{d}\psi(\bar{x})(w) - \langle v_k, w \rangle}{\frac{1}{2}} \geq \infty,$$

which in turn confirms that

$$\varphi^*(v) = \lim_{k \to \infty} \vartheta_{t_k}^*(v_k) = \lim_{k \to \infty} \frac{1}{t_k}\Big(\frac{\psi(\bar{x}) + \psi^*(v_k) - \langle v_k, \bar{x} \rangle}{\frac{1}{2}t_k} + \frac{\mathrm{d}\psi(\bar{x})(w) - \langle v_k, w \rangle}{\frac{1}{2}}\Big) = \infty.$$

If $v \in \partial\psi(\bar{x})$ but $\mathrm{d}\psi(\bar{x})(w) \neq \langle v, w \rangle$, we obtain $\langle v, w \rangle < \mathrm{d}\psi(\bar{x})(w)$. Since we always have

$$\psi(\bar{x}) + \psi^*(v_k) - \langle v_k, \bar{x} \rangle \geq 0 \quad \text{for all } k \in \mathbb{N},$$

we arrive at

$$\varphi^*(v) = \lim_{k \to \infty} \vartheta_{t_k}^*(v_k) \geq \lim_{k \to \infty} \frac{\mathrm{d}\psi(\bar{x})(w) - \langle v_k, w \rangle}{\frac{1}{2}t_k} = \infty,$$

which justifies (4.15) when $v \notin \mathcal{A}(\bar{x}, w)$ and hence finishes the proof. $\qquad\square$

Proposition 4.6 was first established using a different method in [18, Proposition 3.5] for convex piecewise linear-quadratic functions. It was extended in [7, Theorem 3.1] for any convex functions under a restrictive assumption. Indeed, this result demands that the second sub-derivative be the same as the second-order directional derivative. Although this condition holds for convex piecewise linear-quadratic functions, it fails for many important functions occurring in constrained and composite optimization problems including the set indicator functions and eigenvalue functions. As discussed below, however, our assumptions are satisfied for all these examples.

Example 4.7. Suppose that $g : \mathbb{Y} \to \overline{\mathbb{R}}$ is an l.s.c. convex function and $\bar{z} \in \mathbb{Y}$.

(a) If $\mathbb{Y} = \mathbb{R}^m$, g is convex piecewise linear-quadratic (Example 3.2), and $\bar{z} \in \operatorname{dom} g$, then it follows from Example 3.2 and [21, Exercise 13.61] that g is parabolically regular at $\bar{z}$ for every $y \in \partial g(\bar{z})$ and parabolically epi-differentiable at $\bar{z}$ for every $w \in \operatorname{dom} \mathrm{d}g(\bar{z})$, respectively, and thus all the assumptions of Proposition 4.6 are satisfied for this function.

(b) If $\mathbb{Y} = \mathbb{S}^m$, g is either the maximum eigenvalue function $\lambda_{\max}$ from (3.9) or the function σ_i from (3.10), and $A \in \mathbb{S}^n$, then by Example 3.3 g is parabolically regular at A for every $V \in \partial g(A)$. Moreover, we deduce from [23, Proposition 2.2] that g is parabolically epi-differentiable at A for every $W \in \mathbb{S}^n$ and thus all the assumptions of Proposition 4.6 are satisfied for these functions.

(c) If $g = \delta_C$ and $\bar{z} \in C$, where C is a closed convex set in $\mathbb{Y}$ that is parabolically derivable at $\bar{z}$ for every $w \in T_C(\bar{z})$ and parabolically regular at $\bar{z}$ for every $v \in N_C(\bar{z})$, then g satisfies the assumptions imposed in Proposition 4.6. This example of g was recently explored in detail in [14] and encompasses important sets appearing in constrained optimization problems such as polyhedral convex sets, the second-order cone, and the cone of positive semidefinite symmetric matrices.

(d) Assume that g is differentiable at $\bar{z}$ and that there exists a continuous function $h : \mathbb{Y} \to \mathbb{R}$, which is positively homogeneous of degree 2, such that

$$g(z) = g(\bar{z}) + \langle \nabla g(\bar{z}), z - \bar{z} \rangle + \tfrac{1}{2} h(z - \bar{z}) + o(\|z - \bar{z}\|^2).$$

Such a function g is called twice semidifferentiable (cf. [21, Example 13.7]) and often appears in the augmented Lagrangian function associated with (1.1); see [14, Section 8] for more detail. This second-order expansion clearly justifies the parabolic epi-differentiability of g at $\bar{z}$ for every $w \in \mathbb{Y}$. Moreover, one has

$$\mathrm{d}^2 g(\bar{z}, \nabla g(\bar{z}))(w) = h(w) = \mathrm{d}^2 g(\bar{z})(w \mid u) - \langle \nabla g(\bar{z}), u \rangle \quad \text{for all } u, w \in \mathbb{Y},$$

which in turn shows that g is parabolically regular at $\bar{z}$ for $\nabla g(\bar{z})$ due to Proposition 3.6.

It is important to mention that the restrictive assumption on the second subderivative, used in [7, Theorem 3.1], does not hold for cases (b)-(d) in Example 4.7.

5   A Chain Rule for Parabolically Regular Functions

Our main objective in this section is to derive a chain rule for the parabolic regularity of the composite representation (4.3). This opens the door to obtain a chain rule for the second subderivative, and, more importantly, allows us to establish the twice epi-differentiability of the latter composite form.

Taking into account representation (4.3) and picking a subgradient $\bar{v} \in \partial f(\bar{x})$, we define the set of *Lagrangian multipliers* associated with $(\bar{x}, \bar{v})$ by

$$\Lambda(\bar{x}, \bar{v}) = \big\{ y \in \mathbb{Y} \mid \nabla F(\bar{x})^* y = \bar{v}, \ y \in \partial g(F(\bar{x})) \big\}.$$

In what follows, we say that a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ with $(\bar{x}, \bar{v}) \in \mathrm{gph}\, \partial f$ and having the composite representation (4.3) at $\bar{x}$ satisfies the basic assumptions at $(\bar{x}, \bar{v})$ if in addition the following conditions fulfill:

(H1) the metric subregularity constraint qualification holds for the constraint set (4.4) at $\bar{x}$;

(H2) for any $y \in \Lambda(\bar{x}, \bar{v})$, the function g from (4.3) is parabolically epi-differentiable at $F(\bar{x})$ for every $u \in K_g(F(\bar{x}), y)$;

(H3) for any $y \in \Lambda(\bar{x}, \bar{v})$, the function g is parabolically regular at $F(\bar{x})$ for y.

We begin with the following result in which we collect lower and upper estimates for the second subderivative of f taken from (4.3).

Proposition 5.1 (properities of second subderivatives for composite functions). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ have the composite representation (4.3) at $\bar{x} \in \mathrm{dom}\, f$, $\bar{v} \in \partial f(\bar{x})$, and let the basic assumptions (H1) and (H2) hold for f at $(\bar{x}, \bar{v})$. Then the second subderivative $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper l.s.c. function with*

$$\mathrm{dom}\, \mathrm{d}^2 f(\bar{x}, \bar{v}) = K_f(\bar{x}, \bar{v}). \tag{5.1}$$

Moreover, for every $w \in \mathbb{X}$ we have the lower estimate

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) \geq \sup_{y \in \Lambda(\bar{x}, \bar{v})} \left\{ \langle y, \nabla^2 F(\bar{x})(w, w) \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x}) w) \right\}, \tag{5.2}$$

while for every $w \in K_f(\bar{x}, \bar{v})$ we obtain the upper estimate

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) \leq \inf_{z \in \mathbb{X}} \left\{ -\langle z, \bar{v} \rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x}) w \,|\, \nabla F(\bar{x}) z + \nabla^2 F(\bar{x})(w, w)) \right\} < \infty. \tag{5.3}$$

Proof. By Proposition 4.3(ii), we have $\partial^p f(\bar{x}) = \partial f(\bar{x})$. Appealing now to Propositions 2.1(iii) and 3.4 confirms, respectively, that $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper l.s.c. function and that (5.1) holds. The lower estimate (5.2) can be justified as [21, Theorem 13.14] in which this estimate was derived under condition (4.6). To obtain (5.3), observe first that the basic assumption (H1) yields

$$w \in K_f(\bar{x}, \bar{v}) \iff \nabla F(\bar{x}) w \in K_g(F(\bar{x}), y) \tag{5.4}$$

for every $y \in \Lambda(\bar{x}, \bar{v})$. Pick $w \in K_f(\bar{x}, \bar{v})$. Since g is parabolically epi-differentiable at $F(\bar{x})$ for $\nabla F(\bar{x}) w$ due to (H2), Theorem 4.4(iii) implies that f is parabolically epi-differentiable at $\bar{x}$ for w, and so $\mathrm{dom}\, \mathrm{d}^2 f(\bar{x})(w \,|\, \cdot) \neq \emptyset$. This combined with (3.14) and (4.8) results in (5.3) and hence completes the proof. $\square$

While looking simple, the above result carries important information by which we can achieve a chain rule for the second subderivative. To do so, we should look for conditions under which the lower and upper estimates (5.2) and (5.3), respectively, coincide. This motivates us to consider the unconstrained optimization problem

$$\min_{z \in \mathbb{X}} -\langle z, \bar{v} \rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x}) w \,|\, \nabla F(\bar{x}) z + \nabla^2 F(\bar{x})(w, w)). \tag{5.5}$$

When the basic assumptions (H1)-(H3) are satisfied, (5.5) is a convex optimization problem for any $w \in K_f(\bar{x}, \bar{v})$. Using Proposition 4.6 allows us to obtain the dual problem of (5.5) and then examine whether their optimal values coincide. We pursue this goal in the following result.

Theorem 5.2 (duality relationships). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ have the composite representation (4.3) at $\bar{x} \in \mathrm{dom}\, f$, $\bar{v} \in \partial f(\bar{x})$, and let the basic assumptions (H1)-(H3) hold for f at $(\bar{x}, \bar{v})$. Then for each $w \in K_f(\bar{x}, \bar{v})$, the following assertions are satisfied:*

(i) *the dual problem of* (5.5) *is given by*

$$\max_{y\in\mathbb{Y}} \ \langle y, \nabla^2 F(\bar{x})(w,w)\rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \quad \textit{subject to} \ \ y \in \Lambda(\bar{x}, \bar{v}); \qquad (5.6)$$

(ii) *the optimal values of the primal and dual problems* (5.5) *and* (5.6), *respectively, are finite and coincide; moreover, we have* $\Lambda(\bar{x}, \bar{v}, w) \cap (\tau\mathbb{B}) \neq \emptyset$, *where* $\Lambda(\bar{x}, \bar{v}, w)$ *stands for the set of optimal solutions to the dual problem* (5.6) *and where*

$$\tau := \kappa\ell\|\nabla F(\bar{x})\| + \kappa\|\bar{v}\| + \ell \qquad (5.7)$$

with ℓ *and* κ *taken from* (4.3) *and* (4.5), *respectively.*

Proof. Pick $w \in K_f(\bar{x}, \bar{v})$ and observe from (5.4) that $\nabla F(\bar{x})w \in K_g(F(\bar{x}), y)$ for all $y \in \Lambda(\bar{x}, \bar{v})$. This together with Proposition 4.6 ensures that the parabolic subderivative of g at $F(\bar{x})$ for $\nabla F(\bar{x})w$ is a proper, l.s.c., and convex function. Using this combined with [21, Example 11.41] and (4.15) tells us that the dual problem of (5.5) is

$$\max_{y\in\mathbb{Y}} \ \langle y, \nabla^2 F(\bar{x})(w,w)\rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \quad \textit{subject to} \ \ y \in \Lambda(\bar{x}, \bar{v}) \cap \mathcal{D},$$

where $\mathcal{D} := \{y \in \mathbb{Y} \mid \mathrm{d}g(F(\bar{x}))(\nabla F(\bar{x})w) = \langle y, \nabla F(\bar{x})w\rangle\}$. Since $\nabla F(\bar{x})w \in K_g(F(\bar{x}), y)$ for all $y \in \Lambda(\bar{x}, \bar{v})$, we obtain

$$y \in \Lambda(\bar{x}, \bar{v}) \cap \mathcal{D} \iff y \in \Lambda(\bar{x}, \bar{v}).$$

Combining these confirms that the dual problem of (5.5) is equivalent to (5.6) and thus finishes the proof of (i). To prove (ii), consider the optimal value function $\vartheta : \mathbb{Y} \to [-\infty, \infty]$, defined by

$$\vartheta(p) = \inf_{z\in\mathbb{X}} \big\{ -\langle\bar{v}, z\rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \mid \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w,w) + p)\big\}, \quad p \in \mathbb{Y}. \qquad (5.8)$$

We proceed with the following claim:

Claim. *We have* $\partial\vartheta(0) \neq \emptyset$.

To justify the claim, we first need to show $\vartheta(0) \in \mathbb{R}$. To do so, observe that $\bar{v} \in \partial f(\bar{x}) = \partial^p f(\bar{x})$ due to Proposition 4.3(ii). Thus, it follows from Proposition 2.1(iii) and (5.3) that there is a constant $r \in \mathbb{R}_+$ such that for any $w \in K_f(\bar{x}, \bar{v})$ we have

$$-r\|w\|^2 \leq \mathrm{d}^2 f(\bar{x}, \bar{v})(w) \leq \vartheta(0) < \infty,$$

which in turn implies that $\vartheta(0) \in \mathbb{R}$. Next, we are going to show that

$$\vartheta(p) \geq \vartheta(0) - \tau\|p\| \quad \text{for all } p \in \mathbb{X}, \qquad (5.9)$$

where τ is taken from (5.7). To this end, take $(p, z) \in \mathbb{Y} \times \mathbb{X}$ such that

$$u_p := \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w,w) + p \in \mathrm{dom}\,\mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \mid \cdot).$$

By (4.9), we get $u_p \in T^2_{\mathrm{dom}\,g}(F(\bar{x}), \nabla F(\bar{x})w)$. Define now the set-valued mapping $S_w : \mathbb{Y} \rightrightarrows \mathbb{X}$ by

$$S_w(p) := \big\{z \in \mathbb{X} \mid \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w,w) + p \in T^2_{\mathrm{dom}\,g}(F(\bar{x}), \nabla F(\bar{x})w)\big\}, \quad p \in \mathbb{Y}.$$

So, we get $z \in S_w(p)$. It was recently observed in [14, Theorem 4.3] that the mapping S_w enjoys the uniform outer Lipschitzian property at 0 with constant κ taken from (4.5), namely for every $p \in \mathbb{Y}$ we have

$$S_w(p) \subset S_w(0) + \kappa\|p\|\mathbb{B}.$$

22

This combined with $z \in S_w(p)$ results in the existence of $z_0 \in S_w(0)$ and $b \in \mathbb{B}$ such that $z = z_0 + \kappa\|p\|b$. It follows from (4.9) and $z_0 \in S_w(0)$ that

$$\nabla F(\bar{x})z_0 + \nabla^2 F(\bar{x})(w, w) \in \operatorname{dom} \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, \cdot).$$

Since we have

$$u_p - \big(\nabla F(\bar{x})z_0 + \nabla^2 F(\bar{x})(w, w)\big) = p + \kappa\|p\|\nabla F(\bar{x})b,$$

and since the parabolic subderivative $\mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, \cdot)$ is Lipschitz continuous relative to its domain due to Proposition 4.5, we get the relationships

$$
\begin{aligned}
-\langle \bar{v}, z\rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u_p) \;\geq\; & -\langle \bar{v}, z_0\rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, \nabla F(\bar{x})z_0 + \nabla^2 F(\bar{x})(w, w)) \\
& -\ell\, \|p + \kappa\,\|p\|\nabla F(\bar{x})b\,\| - \kappa\|p\|\langle \bar{v}, b\rangle \\
\;\geq\; & \vartheta(0) - \big(\ell\kappa\|\nabla F(\bar{x})\| + \kappa\|\bar{v}\| + \ell\big)\|p\|,
\end{aligned}
$$

which together with (5.7) justify (5.9). Remember that the parabolic subderivative of g at $F(\bar{x})$ for $\nabla F(\bar{x})w$ is a proper and convex function. This implies that the function

$$(z, p) \mapsto -\langle \bar{v}, z\rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w, w) + p)$$

is convex on $\mathbb{X} \times \mathbb{Y}$. Using this together with [21, Proposition 2.22] tells us that ϑ is a convex function on $\mathbb{Y}$. Thus, we conclude from (5.9) and [14, Proposition 5.1] that there exists a subgradient $\bar{y}$ of ϑ at 0 such that

$$\bar{y} \in \partial\vartheta(0) \cap (\tau\mathbb{B}), \tag{5.10}$$

which completes the proof of the claim.

Employing now (5.10) and [2, Theorem 2.142] confirms that the optimal values of the primal and dual problems (5.5) and (5.6), respectively, coincide and that

$$\Lambda(\bar{x}, \bar{v}, w) = \partial\vartheta(0).$$

This together with (5.10) justifies (ii) and hence completes the proof. □

The above theorem extends the recent results obtained in [14, Propositions 5.4 & 5.5] for constraint sets, namely when the function g in (4.3) is the indicator function of a closed convex set. We should add here that for constraint sets, the dual problem (5.6) can be obtained via elementary arguments. However, for the composite form (4.3) a similar result requires using rather advanced theory of epi-convergence.

Remark 5.3 (duality relationship under metric regularity). In the framework of Theorem 5.2, we want to show that replacing assumption (H1) with the strictly stronger constraint qualification (4.6) allows us not only to drop the imposed Lipschitz continuity of g from (4.3) but also to simplify the proof of Theorem 5.2. To this end, let $w \in K_f(\bar{x}, \bar{v})$ and define the function

$$\psi(u) := \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \,|\, u), \quad u \in \mathbb{X}.$$

By Proposition 4.6, ψ is a proper, l.s.c., and convex function. Employing [21, Proposition 13.12] tells us that

$$T_{\operatorname{epi} g}(p) + T^2_{\operatorname{epi} g}(p, q) \subset T^2_{\operatorname{epi} g}(p, q) = \operatorname{epi} \psi, \tag{5.11}$$

where $p := \big(F(\bar{x}), g(F(\bar{x}))\big)$ and $q := \big(\nabla F(\bar{x})w, \mathrm{d}g(F(\bar{x}))(\nabla F(\bar{x})w)\big)$ and where the equality in the right side comes from [21, Example 13.62(b)]. We are going to show that the validity of (4.6) yields

$$N_{\operatorname{dom} \psi}(u) \cap \ker \nabla F(\bar{x})^* = \{0\} \tag{5.12}$$

for any $u \in \operatorname{dom} \psi$. To this end, pick $u \in \operatorname{dom} \psi$ and conclude from (5.11) and [21, Exercise 6.44] that

$$N_{\operatorname{epi}\psi}\big(u, \psi(u)\big) \subset N_{T_{\operatorname{epi}g}(p)}(0) \cap N_{\operatorname{epi}\psi}\big(u, \psi(u)\big) = N_{\operatorname{epi}g}(p) \cap N_{\operatorname{epi}\psi}\big(u, \psi(u)\big).$$

This together with (4.6) and the relationship $N_{\operatorname{dom}\psi}(u) = \partial^\infty \psi(u)$ stemming from the convexity of ψ confirms the validity of (5.12). Appealing now to [2, Theorem 2.165] gives another proof of Theorem 5.2 when assumption (H1) therein is replaced with the strictly stronger constraint qualification (4.6).

The established duality relationships in Theorem 5.2 open the door to derive a chain rule for parabolically regular functions and to find an exact chain rule for the second subderivative of the composite function (4.3) under our basic assumptions.

Theorem 5.4 (chain rule for parabolic regularity). *Let $f : \mathbb{X} \to \overline{\mathbb{R}}$ have the composite representation (4.3) at $\bar{x} \in \operatorname{dom} f$, $\bar{v} \in \partial f(\bar{x})$, and let the basic assumptions (H1)-(H3) hold for f at $(\bar{x}, \bar{v})$. Then f is parabolically regular at $\bar{x}$ for $\bar{v}$. Furthermore, for every $w \in \mathbb{X}$, the second subderivative of f at $\bar{x}$ for $\bar{v}$ is calculated by*

$$
\begin{aligned}
\mathrm{d}^2 f(\bar{x}, \bar{v})(w) &= \max_{y \in \Lambda(\bar{x}, \bar{v})} \big\{ \langle y, \nabla^2 F(\bar{x})(w, w) \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \big\} \quad (5.13) \\
&= \max_{y \in \Lambda(\bar{x}, \bar{v}) \cap (\tau \mathbb{B})} \big\{ \langle y, \nabla^2 F(\bar{x})(w, w) \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \big\},
\end{aligned}
$$

where τ is taken from (5.7).

Proof. It was recently observed in [13, Corollary 3.7] that the Lagrange multiplier set $\Lambda(\bar{x}, \bar{v})$ enjoys the following property:

$$\Lambda(\bar{x}, \bar{v}) \cap (\tau \mathbb{B}) \neq \emptyset. \tag{5.14}$$

Take $w \in K_f(\bar{x}, \bar{v})$. By (5.2) and Theorem 5.2(ii), we obtain

$$\max_{y \in \Lambda(\bar{x}, \bar{v})} \big\{ \langle y, \nabla^2 F(\bar{x})(w, w) \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \big\} \leq \mathrm{d}^2 f(\bar{x}, \bar{v})(w). \tag{5.15}$$

On the other hand, using (3.14), (4.8), and Theorem 5.2(ii), respectively, gives us the inequalities

$$
\begin{aligned}
\mathrm{d}^2 f(\bar{x}, \bar{v})(w) &\leq \inf_{z \in \mathbb{X}} \big\{ \mathrm{d}^2 f(\bar{x})(w \mid z) - \langle z, \bar{v} \rangle \big\} \\
&= \inf_{z \in \mathbb{X}} \big\{ -\langle z, \bar{v} \rangle + \mathrm{d}^2 g(F(\bar{x}))(\nabla F(\bar{x})w \mid \nabla F(\bar{x})z + \nabla^2 F(\bar{x})(w, w)) \big\} \\
&= \max_{y \in \Lambda(\bar{x}, \bar{v}) \cap (\tau \mathbb{B})} \big\{ \langle y, \nabla^2 F(\bar{x})(w, w) \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \big\}.
\end{aligned}
$$

These combined with (5.15) ensure that the claimed second subderivative formulas for f at $\bar{x}$ for $\bar{v}$ hold for any $w \in K_f(\bar{x}, \bar{v})$ and that

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \inf_{z \in \mathbb{X}} \big\{ \mathrm{d}^2 f(\bar{x})(w \mid z) - \langle z, \bar{v} \rangle \big\} \quad \text{for all } w \in K_f(\bar{x}, \bar{v}).$$

Appealing now to Proposition 3.6, we conclude that f is parabolically regular at $\bar{x}$ for $\bar{v}$.

What remains is to validate the second subderivative formulas for $w \notin K_f(\bar{x}, \bar{v})$. It follows from Theorem 4.4(iii) that f is parabolically epi-differentiable at $\bar{x}$ for every $w \in K_f(\bar{x}, \bar{v})$ and thus $\operatorname{dom} \mathrm{d}^2 f(\bar{x})(w \mid \cdot) \neq \emptyset$ for every $w \in K_f(\bar{x}, \bar{v})$. So, by Proposition 3.4 we have $\operatorname{dom} \mathrm{d}^2 f(\bar{x}, \bar{v}) = K_f(\bar{x}, \bar{v})$. Since the second subderivative $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function, we obtain $\mathrm{d}^2 f(\bar{x}, \bar{v})(w) = \infty$ for all $w \notin K_f(\bar{x}, \bar{v})$. On the other hand, we understand from (5.4) that $w \notin K_f(\bar{x}, \bar{v})$ amounts to $\nabla F(\bar{x})w \notin K_g(F(\bar{x}), y)$ for every $y \in \Lambda(\bar{x}, \bar{v})$. Combining the basic assumption (H2) and Proposition 3.4 tells us that for every $y \in \Lambda(\bar{x}, \bar{v})$ we have $\mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) = \infty$ whenever $w \notin K_f(\bar{x}, \bar{v})$. This together with (5.14) confirms that both sides in (5.13) are ∞ for every $w \notin K_f(\bar{x}, \bar{v})$ and thus the claimed formulas for the second subderivative of f hold for this case. This completes the proof. □

A chain rule for parabolic regularity of the composite function (4.3), where g is not necessarily locally Lipschitz continuous relative to its domain, was established in [2, Proposition 3.104]. The assumptions utilized in the latter result were stronger than those used in Theorem 5.4. Indeed, [2, Proposition 3.104] assumes that g is second-order regular in the sense of [2, Definition 3.93] and the basic constraint qualification (4.6) is satisfied and uses a different approach to derive this result. When g is a convex piecewise linear-quadratic, parabolic regularity of the composite function (4.3) was established in [21, Theorem 13.67] under the stronger condition (4.6). Theorem 5.4 covers the aforementioned results and shows that we can achieve a similar conclusion under the significantly weaker condition (4.5).

As an immediate consequence of the above theorem, we can easily guarantee the twice epi-differentiability of the composite form (4.3) under our basic assumptions.

Corollary 5.5 (chain rule for twice epi-differentiability). *Let the function f from (4.3) satisfy all the assumptions of Theorem 5.4. Then f is twice epi-differentiable at $\bar{x}$ for $\bar{v}$.*

Proof. By Theorem 4.4(iii), f is parabolically epi-differentiable at $\bar{x}$ for every $w \in K_f(\bar{x}, \bar{v})$. Employing now Theorems 5.4 and 3.8 implies that f is twice epi-differentiable at $\bar{x}$ for $\bar{v}$. $\qquad\square$

Remark 5.6 (discussion on twice epi-differentiability). Corollary 5.5 provides a far-going extension of the available results for the twice epi-differentiability of extended-real-valued functions. To elaborate more, suppose that $f : \mathbb{X} \to \overline{\mathbb{R}}$ has a composite form (4.3) at $\bar{x} \in \mathrm{dom}\, f$. Then the following observations hold:

(a) If $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{R}^m$, and g in (4.3) is convex piecewise linear-quadratic, then Rockafellar proved in [18] that under the fulfillment of the basic constraint qualification (4.6), f is twice epi-differentiable. This result was improved recently in [13, Theorem 5.2], where it was shown that using the strictly weaker condition (4.5) in the Rockafellar's framework [18] suffices to ensure the twice epi-differentiability of f. Taking into account Example 4.7(a) tells us both these results can be derived from Corollary 5.5.

(b) If $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{S}^m$, and g is either the maximum eigenvalue function $\lambda_{\max}$ from (3.9) or the function σ_i from (3.10), then we fall into the framework considered by Turki in [24, Theorems 2.3 & 2.5] in which he justified the twice epi-differentiability of f. Since in this framework we have $\mathrm{dom}\, g = \mathbb{S}^m$, both conditions (4.6) and (4.5) are automatically satisfied. By Example 4.7(b), the twice epi-differentiability of f can be deduced from Corollary 5.5.

(c) If $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{R}^m$, and $g = \delta_C$ with the closed convex set C taken from Example 4.7(c), we fall into the framework considered in [14]. In this case, Corollary 5.5 can cover the twice epi-differentiability of f obtained in [14, Corollary 5.11].

(d) If $\mathbb{X} = \mathbb{R}^n$, $\mathbb{Y} = \mathbb{R}^m$, and g is a proper, convex, l.s.c., and positively homogeneous, then we fall into the framework, considered by Shapiro in [22]. In this case, the composite form (4.3) is called *decomposable*; see [9, 22] for more detail about this class of extended-real-valued functions. It was proven in [9, Lemma 5.3.27] that for this case of g, the composite form (4.3) is twice epi-differentiable if it is convex and if the nondegeneracy condition for this setting holds; see [9, Definition 5.3.1] for the definition of this condition. In this framework, by the positive homogeneity of g and $F(\bar{x}) = 0$, coming from [9, Definition 5.3.1], we can easily show that g is parabolically regular. Moreover the assumed nondegeneracy condition in [9, Lemma 5.3.27] yields the validity of condition (4.6). As pointed out in Remark 5.3, the Lipschitz continuity of g in the composite form (4.3) can be relaxed when condition (4.6) is satisfied. Since the nondegeneracy condition implies that the set of Lagrange multipliers $\Lambda(\bar{x}, \bar{v})$ is a singleton, we can use estimates (5.2) and (5.3) to justify parabolic regularity of the composite form (4.3) in the framework of [9]. This together

25

with Corollary 5.5 allows to recover [9, Lemma 5.3.27]. Furthermore, we can drop the convexity of the composite form (4.3), assumed in [9].

6 Second-Order Optimality Conditions for Composite Problems

In this section, we focus mainly on obtaining second-order optimality conditions for the composite problem (1.1), where $\varphi : \mathbb{X} \to \mathbb{R}$ and $F : \mathbb{X} \to \mathbb{Y}$ are twice differentiable and the function $g : \mathbb{Y} \to \overline{\mathbb{R}}$ is an l.s.c. convex function that is locally Lipschitz continuous relative to its domain. The latter means that for any $y \in \mathrm{dom}\, g$, the function g is Lipschitz continuous around y relative to its domain. Important examples of constrained and composite optimization problems can be achieved when g is one of the functions considered in Example 4.7. For any pair $(x, y) \in \mathbb{X} \times \mathbb{Y}$, the Lagrangian associated with the composite problem (1.1) is defined by

$$L(x, y) = \varphi(x) + \langle F(x), y \rangle - g^*(y),$$

where g^* is the Fenchel conjugate of the convex function g. We begin with the following result in which we collect second-order optimality conditions for (1.1) when our basic assumptions are satisfied. Recall that a point $\bar{x} \in \mathbb{X}$ is called a feasible solution to the composite problem (1.1) if we have $F(\bar{x}) \in \mathrm{dom}\, g$.

Theorem 6.1 (second-order optimality conditions). *Let $\bar{x}$ be a feasible solution to problem (1.1) and let $f := g \circ F$ and $\bar{v} := -\nabla\varphi(\bar{x}) \in \partial f(\bar{x})$ with φ, g, and F taken from (1.1). Assume that the basic assumptions* (H1)-(H3) *hold for f at $(\bar{x}, \bar{v})$. Then the following second-order optimality conditions for the composite problem (1.1) are satisfied:*

(i) *if $\bar{x}$ is a local minimum of (1.1), then the second-order necessary condition*

$$\max_{y \in \Lambda(\bar{x}, \bar{v})} \left\{ \langle \nabla^2_{xx} L(\bar{x}, y) w, w \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \right\} \geq 0$$

holds for all $w \in K_f(\bar{x}, \bar{v})$;

(ii) *the validity of the second-order condition*

$$\max_{y \in \Lambda(\bar{x}, \bar{v})} \left\{ \langle \nabla^2_{xx} L(\bar{x}, y) w, w \rangle + \mathrm{d}^2 g(F(\bar{x}), y)(\nabla F(\bar{x})w) \right\} > 0 \quad \text{for all} \quad w \in K_f(\bar{x}, \bar{v}) \setminus \{0\}$$

(6.1)

amounts to the existence of constants $\ell > 0$ and $\varepsilon > 0$ such that the second-order growth condition

$$\psi(x) \geq \psi(\bar{x}) + \tfrac{\ell}{2} \|x - \bar{x}\|^2 \quad \text{for all} \ x \in \mathbb{B}_\varepsilon(\bar{x})$$

(6.2)

holds, where $\psi := \varphi + g \circ F$.

Proof. To justify (i), note that since $\bar{x}$ is a local minimum of (1.1), it is a local minimum of $\psi = \varphi + f$. Moreover, $-\nabla\varphi(\bar{x}) \in \partial f(\bar{x})$ amounts to $0 \in \partial\psi(\bar{x})$. Thus, by definition, we arrive at $\mathrm{d}^2\psi(\bar{x}, 0)(w) \geq 0$ for all $w \in \mathbb{X}$. Since φ is twice differentiable at $\bar{x}$, we obtain the following sum rule for the second subderivatives:

$$\mathrm{d}^2\psi(\bar{x}, 0)(w) = \langle \nabla^2\varphi(\bar{x})w, w \rangle + \mathrm{d}^2 f(\bar{x}, \bar{v})(w) \quad \text{for all} \ w \in \mathbb{X}.$$

(6.3)

Combing these with the chain rule (5.13) proves (i).

Turing now to (ii), we infer from [21, Theorem 13.24(c)] that $\mathrm{d}^2\psi(\bar{x}, 0)(w) > 0$ for all $w \in \mathbb{X} \setminus \{0\}$ amounts to the existence of some constants $\ell > 0$ and $\varepsilon > 0$ for which the second-order growth condition (6.2) holds. Remember from (5.1) and (6.3) that

$$\mathrm{dom}\, \mathrm{d}^2\psi(\bar{x}, 0) = \mathrm{dom}\, \mathrm{d}^2 f(\bar{x}, \bar{v}) = K_f(\bar{x}, \bar{v}).$$

(6.4)

Using these, the chain rule (5.13), and the sum rule (6.3) proves the claimed equivalence in (ii) and thus finishes the proof. □

Remark 6.2 (discussion on second-order optimality conditions). The second-order optimality conditions for composite problems were established in [2, Theorems 3.108 & 3.109] for (1.1) by expressing (1.1) equivalently as a constrained problem and then appealing to the theory of second-order optimality conditions for the latter class of problems. While not assuming that g is locally Lipschitz continuous relative to its domain, theses results were established under condition (4.6) and the second-order regularity in the sense of [2, Definition 3.93] that are strictly stronger than condition (4.5) and the parabolic regularity, respectively, we imposed in Theorem 6.1. Another major difference is that we require that g be parabolically epi-differentiable (assumption (H2)), which was not assumed in [2]. This assumption plays an important role in our developments and has two important consequences: 1) it makes the parabolic subderivative be a *convex* function and help us obtain a precise formula for the Fenchel conjugate of the parabolic subderivative in our framework; 2) it allows to establish the equivalence between (6.1) and the growth condition in Theorem 6.1. These facts were not achieved in [2]; indeed, [2, Theorem 3.109] was written in terms of the conjugate of the parabolic subderivative and only states that condition (6.1) implies the growth condition therein.As discussed in Remark 5.3, if we replace condition (4.5) with the stronger condition (4.6), the imposed Lipschitz continuity of g can be relaxed in our developments. It is worth mentioning that the imposed Lipschitz continuity of g relative to its domain, utilized in this paper, does not seem to be restrictive and allows us to provide an umbrella under which second-order variational analysis for composite problems can be carried out under condition (4.6) in the same level of perfection as those for constrained problems. We believe that if we strengthen condition (4.6) to the metric subregularity of the epigraphical mapping $(x, \alpha) \mapsto (F(x), \alpha) - \text{epi } g$, the imposed Lipschitz continuity of g can be relaxed in our developments.

Cominetti [7, Theorem 5.1] established second-order optimality conditions for the composite problem (1.1) similar to Theorem 6.1 without making a connection between (6.1) and the growth condition (6.2). As mentioned in our discussion after Example 4.7, the results in [7] were established under condition (4.6) and a restrictive assumption on the second subderivative, which does not hold for important classes of composite problems. When we are in the framework of Remark 5.6(a), Theorem 6.1 was first achieved by Rockafellar in [19, Theorem 4.2] under condition (4.6) and was improved recently in [13, Theorem 6.2] by replacing the latter condition with (4.5). For the framework of Remark 5.6(b), the second-order optimality conditions from Theorem 6.1 were obtained in [24, Theorem 4.2]. Finally, if we are in the framework of Remark 5.6(c), Theorem 6.1 covers our recent developments in [14].

We end this section by obtaining a characterization of strong metric subregularity of the subgradient mapping of the objective function of the composite problem (1.1).

Theorem 6.3 (strong metric subregularity of the subgradient mappings in composite problems). *Let $\bar{x}$ be a feasible solution to problem (1.1) and let $f := g \circ F$ and $\bar{v} := -\nabla\varphi(\bar{x}) \in \partial f(\bar{x})$ with φ, g, and F taken from (1.1). Assume that the basic assumptions (H1)-(H3) hold for f at $(\bar{x}, \bar{v})$ and that both φ and F are $\mathcal{C}^2$-smooth around $\bar{x}$. Then the following conditions are equivalent:*

  (i) *the point $\bar{x}$ is a local minimizer for $\psi = \varphi + f$ and the subgradient mapping $\partial\psi$ is strongly metrically subregular at $(\bar{x}, 0)$;*

  (ii) *the second-order sufficient condition (6.1) holds.*

Proof. We conclude from (6.3) and (6.4) that (6.1) amounts to the fulfillment of the condition

$$\mathrm{d}^2\psi(\bar{x}, 0)(w) > 0 \quad \text{for all } w \in \mathbb{X} \setminus \{0\}. \tag{6.5}$$

If (i) holds, we conclude from the local optimality of $\bar{x}$ that $\mathrm{d}^2\psi(\bar{x}, 0)(w) \geq 0$ for all $w \in \mathbb{X}$. Since (ii) is equivalent to (6.5), it suffices to show that there is no $w \in \mathbb{X} \setminus \{0\}$ such that

$\mathrm{d}^2\psi(\bar{x},0)(w) = 0$. Suppose on the contrary that there exists $\bar{w} \in \mathbb{X} \setminus \{0\}$ satisfying the latter condition. This means that $\bar{w}$ is a minimizer for the problem

$$\text{minimize } \tfrac{1}{2}\mathrm{d}^2\psi(\bar{x},0)(w) \quad \text{subject to} \quad w \in \mathbb{X}.$$

Since both φ and F are $\mathcal{C}^2$-smooth around $\bar{x}$, we can show using similar arguments as [13, Proposition 7.1] that ψ is prox-regular and subdifferentially continuousat $\bar{x}$ for 0. This together with the Fermat stationary principle and (3.21) results in

$$0 \in \partial\big(\tfrac{1}{2}\mathrm{d}^2\psi(\bar{x},0)\big)(\bar{w}) = D(\partial\psi)(\bar{x},0)(\bar{w}). \tag{6.6}$$

Since $\partial\psi$ is strongly metrically subregular at $(\bar{x},0)$, we deduce from (2.2) that $\bar{w} = 0$, a contradiction. This proves (ii).

To justify the opposite implication, assume that (ii) holds. According to Theorem 6.1(ii), $\bar{x}$ is a local minimizer for ψ. Pick now $w \in \mathbb{X}$ such that $0 \in D(\partial\psi)(\bar{x},0)(w)$. To obtain (i), we require by (2.2) to show that $w = 0$. Employing now (6.6) yields $0 \in \partial\big(\tfrac{1}{2}\mathrm{d}^2\psi(\bar{x},0)\big)(w)$. This combined with [6, Lemma 3.7] confirms that $\mathrm{d}^2\psi(\bar{x},0)(w) = \langle 0, w \rangle = 0$. Remember that (ii) is equivalent to (6.5). Combining these results in $w = 0$ and thus proves (i). $\square$

The above result was first observed in [8, Theorem 4G.1] for a subclass of nonlinear programming problems and was extended in [6, Theorem 4.6] for $\mathcal{C}^2$-cone reducible constrained optimization problems and in [14, Theorem 9.2] for parabolically regular constrained optimization problems. The theory of the twice epi-differentiability, obtained in this paper, provides an easy path to achieve a similar result for the composite problem (1.1).

It is worth mentioning that similar characterizations as [14, Theorem 4.2] can be achieved for the KKT system of (1.1). Furthermore, Corollary 3.9 provides a systematic method to calculate proto-derivatives of subgradient mappings of functions enjoying the composite form (4.3), a path we will pursue in our future research.

References

[1] A. BEN-TAL AND J. ZOWE, *A unified theory of first- and second-order conditions for extremum problems in topological vector spaces*, Math. Program. 19 (1982), 39–76.

[2] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[3] J. F. BONNANS, R. COMINETTI AND A. SHAPIRO, *Second-order optimality conditions based on parabolic second-order tangent sets*, SIAM J. Optim. 9 (1998), 466–492.

[4] J. V. BURKE, A. S. LEWIS, AND M. L. OVERTON, *Optimal stability and eigenvalue multiplicity*, Found. Comput. Math. 1 (2001), 205–225.

[5] R. W. CHANEY, *On second derivatives for nonsmooth functions*, Nonlinear Anal. 9 (1985), 1189–1209.

[6] N. H. CHIEU, L. V. HIEN, T. T. A. NGHIA, AND H. A. TUAN, *Second-order optimality conditions for strong local minimizers via subgradient graphical derivative*, (2019), arXiv:1903.05746.

[7] R. COMINETTI, *On pseudo-differentiability*, Trans. Amer. Math. Soc. 324 (1991), 843–847.

[8] A. L. DONTCHEV AND R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd edition, Springer, Dordrecht, 2014.

[9] A. MILZAREK, *Numerical Methods and Second Order Theory for Nonsmooth Problems*, Ph.D. Dissertation, University of Munich, 2016.

[10] A.D. IOFFE, *Variational analysis of a composite function: A formula for the lower second-order epi-derivative*, J. Math. Anal. Appl. 160 (1991) 379–405.

[11] A. B. LEVY, *Second-order epi-derivatives of composite functionals*, Ann. Oper. Res. 101 (2001), 267–281.

[12] B. S. MORDUKHOVICH, *Variational Analysis and Applications*, Springer Monographs in Mathematics, Springer, Cham, Switzerland, 2018.

[13] A. MOHAMMADI, B. MORDUKHOVICH AND M. E. SARABI, *Variational analysis of composite models with applications to continuous optimization*, To appear in Math. Oper. Res. (2020), arXiv:1905.08837.

[14] A. MOHAMMADI, B. MORDUKHOVICH AND M. E. SARABI, *Parabolic regularity via geometric variational analysis*, (2019), arXiv:1909.00241.

[15] A. MOHAMMADI, B. MORDUKHOVICH AND M. E. SARABI, *Stability of KKT systems and superlinear convergence of the SQP method under parabolic regularity*, (2019), arXiv:1910.06894.

[16] R. A. POLIQUIN AND R. T. ROCKAFELLAR, *Prox-regular functions in variational analysis* Trans. Amer. Math. Soc. 348 (1996), 1805–1838.

[17] R. T. ROCKAFELLAR, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. H. Poincaré Analyse Non Linéaire 2 (1985), 167–184.

[18] R. T. ROCKAFELLAR, *first- and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc. 307 (1988), 75–108.

[19] R. T. ROCKAFELLAR, *Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives*, Math. Oper. Res. 14 (1989), 462–484.

[20] R. T. ROCKAFELLAR, *Generalized second derivatives of convex functions and saddle functions*, Trans. Amer. Math. Soc. 322 (1990), 51–77.

[21] R. T. ROCKAFELLAR AND R. J-B. WETS, *Variational Analysis*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 317, Springer, Berlin, 2006.

[22] A. SHAPIRO, *On a class of nonsmooth composite functions*, Math. Oper. Res. 28 (2003), 677–692.

[23] M. TORKI, *Second-order directional derivatives of all eigenvalues of a symmetric matrix*, Nonlinear Anal. 46 (2001), 1133–1150.

[24] M. TORKI, *First- and second-order epi-differentiability in eigenvalue optimization*, J. Math. Anal. Appl. 234 (1999), 391–416.