informs
https://pubsonline.informs.org/journal/moor
MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–30
ISSN 0364-765X (print), ISSN 1526-5471 (online)

Parabolic Regularity of Spectral Functions

Ashkan Mohammadi,[a] Ebrahim Sarabi[b,*]

[a] Department of Mathematics, Georgetown University, Washington, District of Columbia 20057; [b] Department of Mathematics, Miami University, Oxford, Ohio 45056

*Corresponding author

Contact: ashkan.math@gmail.com (AM); sarabim@miamioh.edu, https://orcid.org/0000-0001-6049-8902 (ES)

Received: January 10, 2023
Revised: December 21, 2023; April 11, 2024
Accepted: June 28, 2024
Published Online in Articles in Advance: August 12, 2024

MSC2020 Subject Classifications: Primary: 49J52, 15A18, 49J53

https://doi.org/10.1287/moor.2023.0010

Abstract. This paper is devoted to the study of the second-order variational analysis of spectral functions. It is well-known that spectral functions can be expressed as a composite function of symmetric functions and eigenvalue functions. We establish several second-order properties of spectral functions when their associated symmetric functions enjoy these properties. Our main attention is given to characterize parabolic regularity for this class of functions. It was observed recently that parabolic regularity can play a central rule in ensuring the validity of important second-order variational properties, such as twice epi-differentiability. We demonstrates that for convex spectral functions, their parabolic regularity amounts to that of their symmetric functions. As an important consequence, we calculate the second subderivative of convex spectral functions, which allows us to establish second-order optimality conditions for a class of matrix optimization problems.

Funding: The research of A. Mohammadi is funded by a postdoctoral fellowship from Georgetown University. E. Sarabi is partially supported by the U.S. National Science Foundation [Grant DMS 2108546].

Keywords: spectral functions • parabolic regularity • twice epi-differentiability • composite optimization • second-order optimality conditions

1. Introduction

This paper aims to study second-order variational properties, including parabolic regularity and twice epi-differentiability, of spectral functions. These are functions $g : \mathbf{S}^n \to \overline{\mathbf{R}} := [-\infty, \infty]$, where $\mathbf{S}^n$ stands for the real vector space of $n \times n$ symmetric matrices, that are orthogonally invariant; namely, for any $n \times n$ symmetric matrix X and any $n \times n$ orthogonal matrix U, we have

$$g(X) = g(U^\top XU).$$

It is well-known (cf. Lewis [18, proposition 4]) that any spectral function g can be equivalently expressed in a composite form

$$g(X) = (\theta \circ \lambda)(X), \ \ X \in \mathbf{S}^n, \tag{1}$$

where $\theta : \mathbf{R}^n \to \overline{\mathbf{R}}$ is a permutation-invariant function on $\mathbf{R}^n$, called symmetric, and λ is a function, which assigns to each matrix $X \in \mathbf{S}^n$ its eigenvalue vector $(\lambda_1(X), \ldots, \lambda_n(X))$ arranged in nonincreasing order.

Davis [9] showed that convexity of the permutation-invariant function θ in (1) is inherited by the spectral function g. A similar observation was made by Lewis [17] about differentiability and strict differentiability and by Lewis and Sendov [19] about twice differentiability. It was shown in Lewis [17], Lewis [18], and Daniilidis et al. [8], respectively, that the calculation of different notions of subdifferentials of spectral functions and prox-regularity, which plays an important role in second-order variational analysis, enjoys this striking pattern as well.

The main question that we are trying to answer in this paper is whether such a striking pattern can be extended for other important second-order variational properties, including parabolic regularity (see Definition 1) and twice epi-differentiability. Although the former was first introduced more than two decades ago in Rockafellar and Wets [25, definition 13.69], its persuasive role in second-order variational analysis was revealed quite recently in Mohammadi et al. [21] and Mohammadi and Sarabi [20], where it was shown for the first time that any parabolic regular function is twice epi-differentiable at any points in the graph of its subgradient mapping. This observation provided a systematic approach for the study of twice epi-differentiability of extended-real-valued functions, which has important applications in understanding various second-order variational concepts, such as proto-differentiability of subgradient mappings (cf. Mohammadi and Sarabi [20, corollary 3.9]), twice epi-differentiability of the augmented Lagrangian functions associated with composite and constrained optimization problems (cf. Mohammadi et al. [21, theorem 8.3]), and the characterization of the quadratic growth condition for this class of functions (cf. Hang and Sarabi [16, theorem 4.1]). Moreover, parabolic regularity was utilized in Mohammadi et al. [21] and Mohammadi and Sarabi [20] to obtain the *exact* chain rule for the second subderivative of certain composite functions, commonly seen in different classes optimization problems. Such a chain rule has an important application in finding second-order necessary and sufficient optimality conditions for optimization problems.

It was demonstrated in Mohammadi and Sarabi [20] that important second-order variational properties of a composite function $\psi \circ F$, where $\psi : \mathbf{X} \to (-\infty, \infty]$ is convex and $F : \mathbf{X} \to \mathbf{Y}$ is twice differentiable with $\mathbf{X}$ and $\mathbf{Y}$ being finite-dimensional Hilbert spaces, can be established at any $\overline{x} \in \text{dom}\,(\psi \circ F)$, provided that ψ is parabolically regular and that the following metric subregularity constraint qualification is satisfied (cf. Mohammadi and Sarabi [20, definition 4.2]): There exists a constant $\kappa \geq 0$, such that the estimate

$$\text{dist}(x,\ \text{dom}\,(\psi \circ F)) \leq \kappa\, \text{dist}(F(x),\ \text{dom}\,\psi), \tag{2}$$

holds for all x sufficiently close to $\overline{x}$. Thus, it is natural to ask whether a similar approach can be utilized for the composite representation in (1) of spectral functions. To do so, two major obstacles seem to hinder proceeding with the approach in Mohammadi and Sarabi [20]: (1) the lack of twice differentiability of the inner function $\lambda(\cdot)$ in (1); and (2) the validity of a constraint qualification similar to the aforementioned condition for the Composite Form (1). Given the Composite Representation (1), it follows from Daniilidis et al. [8, proposition 2.3] that for any $X \in \mathbf{S}^n$, the equality

$$\text{dist}(X,\ \text{dom}\, g) = \text{dist}(\lambda(X),\ \text{dom}\,\theta), \tag{3}$$

always holds. This simple, but important, observation from Daniilidis et al. [8] tells us that the required constraint qualification for dealing with the Composite Form (1) is automatically satisfied. Moreover, looking closer into the established theory in Mohammadi et al. [21] and Mohammadi and Sarabi [20] tells us that twice differentiability of the inner function was not required. Indeed, a quadratic expansion will suffice to proceed in both these publications. Such a quadratic expansion, which is of a parabolic type, is already achieved in Torki [27] for eigenvalue functions. These open the door for using the approach from Mohammadi et al. [21] and Mohammadi and Sarabi [20] to study second-order variational properties of spectral functions.

The outline of the paper is as follows. Section 2 recalls important notation and concepts related to the eigenvalue function. In Section 3, we begin with establishing a chain rule for subderivative of the spectral function in (1). Section 4 is devoted to the study of the parabolic drivability of spectral sets. We will obtain a chain rule for the parabolic subderivative, which plays a central role in the study of parabolic regularity of spectral functions. It is also shown that the parabolic subderivative is a symmetric function with respect to a subset of the space of orthogonal matrices. In Section 5, we demonstrate that the spectral function g in (1) is parabolically regular if and only if the symmetric function θ in (1) enjoys this property. As a consequence, we are going to calculate the second subderivative of spectral functions when the symmetric functions associated with them are convex. This allows us to find second-order optimality conditions for a class of matrix optimization problems.

2. Notation

In what follows, $\mathbf{X}$ and $\mathbf{Y}$ are finite-dimensional Hilbert spaces. By $\mathbf{B}$, we denote the closed unit ball in the space in question, and by $\mathbf{B}_r(x) := x + r\mathbf{B}$, the closed ball centered at x with radius $r > 0$. For any set $C \subset \mathbf{X}$, its indicator function is defined by $\delta_C(x) = 0$ for $x \in C$ and $\delta_C(x) = \infty$ otherwise. We denote by $\text{dist}(x, C)$ the distance between $x \in \mathbf{X}$ and a set C. For $v \in \mathbf{X}$, the subspace $\{w \in \mathbf{X} \mid \langle w, v\rangle = 0\}$ is denoted by $[v]^\perp$. We denote by $\mathbf{R}_+$ (respectively, $\mathbf{R}_-$) the set of nonnegative (respectively, nonpositive) real numbers. Given an $n \times n$ matrix Z and index sets $I, J \subseteq \{1, \ldots, n\}$, denote by Z_{IJ} the submatrix of Z obtained by removing all the rows of Z not in I and all the columns of Z not in J. The matrix Z_I is the submatrix of Z with columns specified by I. Particularly, Z_i is the i-th column of Z, and Z_{ij} is the entry of Z at (i, j) position. Denote by $Z^\dagger$ the Moore-Penrose generalized inverse of Z. Finally, the cardinality of the set $I \subset \mathbf{N}$, where $\mathbf{N}$ stands for the set of natural numbers, is denoted by $|I|$.

Throughout this paper, we denote by $\mathbf{R}^{n\times m}$ the space of all real $n \times m$ matrices and by $\mathbf{S}^n$ the space of all real $n \times n$ symmetric matrices equipped with the inner product

$$\langle X, Y\rangle = \text{tr}\,(XY), \quad X, Y \in \mathbf{S}^n.$$

The induced Frobenius norm of $X \in \mathbf{S}^n$ is defined via the trace inner product by $\|X\| = \sqrt{\text{tr}\,(X^2)}$. Given $X \in \mathbf{S}^n$, its eigenvalues, in nonincreasing order, are denoted by

$$\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X).$$

For any vector $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, denote by Diag (x), the diagonal matrix whose i-th diagonal entry is x_i for any $i = 1, \ldots, n$. The set of all real $n \times n$ orthogonal matrices is denoted by $\mathbf{O}^n$. It is known that for any $X \in \mathbf{S}^n$, there exists an orthogonal matrix U, for which we have

$$X = U \text{ Diag } (\lambda(X))U^\top \quad \text{with} \quad \lambda(X) := (\lambda_1(X), \ldots, \lambda_n(X)). \tag{4}$$

For a given matrix $X \in \mathbf{S}^n$, the set of such orthogonal matrices U is denoted by $\mathbf{O}^n(X)$. We say that two matrices $X, Y \in \mathbf{S}^n$ admit a simultaneous spectral decomposition if there exists $U \in \mathbf{O}^n$ such that $U^\top XU$ and $U^\top YU$ are diagonal matrices. The matrices X and Y are said to have a simultaneous *ordered* spectral decomposition if there exists $U \in \mathbf{O}^n$ such that $U^\top XU = \text{Diag } (\lambda(X))$ and $U^\top YU = \text{Diag } (\lambda(Y))$. It is well-known that for any two matrices $X, Y \in \mathbf{S}^n$, the estimate

$$\|\lambda(X) - \lambda(Y)\| \le \|X - Y\|, \tag{5}$$

always holds. Moreover, equality in this estimate amounts to X and Y admitting a simultaneous ordered spectral decomposition. It is not hard to see that the estimate in (5) amounts to the trace inequality, known as Fan's inequality,

$$\langle X, Y \rangle \le \langle \lambda(X), \lambda(Y) \rangle, \quad X, Y \in \mathbf{S}^n. \tag{6}$$

Assume that $\mu_1(X) > \cdots > \mu_r(X)$ are distinct eigenvalues of $X \in \mathbf{S}^n$ and define then the index sets

$$\alpha_m := \{i \in \{1, \ldots, n\} \,|\, \lambda_i(X) = \mu_m(X)\} \quad \text{for all} \;\; m = 1, \ldots, r. \tag{7}$$

Moreover, define $\ell_i(X)$ for any $i \in \{1, \ldots, n\}$ to be the number of eigenvalues of X that are equal to $\lambda_i(X)$, but are ranked before $\lambda_i(X)$, including $\lambda_i(X)$. This integer allows us to locate $\lambda_i(X)$ in the group of the eigenvalues of X as follows:

$$\lambda_1(X) \ge \cdots \ge \lambda_{i-\ell_i(X)} > \lambda_{i-\ell_i(X)+1}(X) = \cdots = \lambda_i(X) \ge \cdots \ge \lambda_n(X). \tag{8}$$

Note that the index sets α_m present a partition of $\{1, \ldots, n\}$, meaning that $\{1, \ldots, n\} = \cup_{m=1}^r \alpha_m$. In what follows, we often drop X from $\ell_i(X)$ when the dependence of ℓ_i on X can be seen clearly from the context. Given an $n \times n$ matrix W, it is not hard to see that for any $U \in \mathbf{O}^n$ and any $m = 1, \ldots, r$, we always have

$$U_{\alpha_m}^\top UWU^\top U_{\alpha_m} = W_{\alpha_m \alpha_m}. \tag{9}$$

This simple observation will often be utilized in Section 5.

The following estimates are an easy consequence of Torki [27, proposition 1.4] (cf. see the proof of Torki [27, theorem 1.5]) and play a major role in our second-order variational analysis of eigenvalue functions in this paper.

Proposition 1 (First-Order Expansion of Eigenvalue Functions). *Assume that $X \in \mathbf{S}^n$ has the Eigenvalue Decomposition (4) for some $U \in \mathbf{O}^n(X)$. Let $\mu_1 > \cdots > \mu_r$ be distinct eigenvalues of X. Then, for any $H \in \mathbf{S}^n$ that $H \to 0$ and any $i \in \{1, \ldots, n\}$, the estimates*

$$\lambda_i(X + H) = \lambda_i(X) + \lambda_{\ell_i}(U_{\alpha_m}^\top HU_{\alpha_m} + U_{\alpha_m}^\top H(\mu_m I - X)^\dagger HU_{\alpha_m}) + O(\|H\|^3), \tag{10}$$

and

$$\lambda_i(X + H) = \lambda_i(X) + \lambda_{\ell_i}(U_{\alpha_m}^\top HU_{\alpha_m}) + O(\|H\|^2), \tag{11}$$

hold, where $m \in \{1, \ldots, r\}$ with $i \in \alpha_m$.

Note that the Estimate (11) clearly tells us that the eigenvalue function $\lambda_i(\cdot)$, $i \in \{1, \ldots, n\}$, is directionally differentiable at X at any direction $H \in \mathbf{S}^n$, and its directional derivative $\lambda_i'(X; H)$ can be calculated by

$$\lambda_i'(X; H) := \lim_{t \downarrow 0} \frac{\lambda_i(X + tH) - \lambda_i(X)}{t} = \lambda_{\ell_i}(U_{\alpha_m}^\top HU_{\alpha_m}), \tag{12}$$

where $m \in \{1, \ldots, r\}$ with $i \in \alpha_m$. In other words, we have

$$\lambda'(X; H) = (\lambda(U_{\alpha_1}^\top HU_{\alpha_1}), \ldots, \lambda(U_{\alpha_r}^\top HU_{\alpha_r})),$$

where $\lambda(U_{\alpha_m}^\top HU_{\alpha_m}) \in \mathbf{R}^{|\alpha_m|}$ for any $m = 1, \ldots, r$. This observation indicates that both estimates in (10) and (11) are, indeed, a first-order estimate of eigenvalue function $\lambda_i(\cdot)$. To obtain a second-order estimate, we need to repeat a similar argument for each of the symmetric matrices $U_{\alpha_m}^\top HU_{\alpha_m}$ for any $m \in \{1, \ldots, r\}$. To this end, fix $m \in \{1, \ldots, r\}$

and observe that $U_{\alpha_m}^\top HU_{\alpha_m} \in \mathbf{S}^{|\alpha_m|}$. Thus, we find $Q_m \in \mathbf{O}^{|\alpha_m|}(U_{\alpha_m}^\top HU_{\alpha_m})$ such that

$$U_{\alpha_m}^\top HU_{\alpha_m} = Q_m\Lambda(U_{\alpha_m}^\top HU_{\alpha_m})Q_m^\top. \tag{13}$$

Denote by $\eta_1^m > \cdots > \eta_{\rho_m}^m$ the distinct eigenvalues of $U_{\alpha_m}^\top HU_{\alpha_m}$. Similar to (7), define the index sets

$$\beta_j^m := \{i \in \{1,\ldots,|\alpha_m|\} \,|\, \lambda_i(U_{\alpha_m}^\top HU_{\alpha_m}) = \eta_j^m\} \quad \text{for all } j = 1,\ldots,\rho_m. \tag{14}$$

To state the promised second-order estimate for eigenvalue functions, we need to clarify some of the indices, appeared therein. To do so, pick $i \in \{1,\ldots,n\}$, and observe that there is $m \in \{1,\ldots,r\}$ such that $i \in \alpha_m$ and that $\ell_i(X) \in \{1,\ldots,|\alpha_m|\}$, where $\ell_i(X)$ is defined by (8). Furthermore, we find $j \in \{1,\ldots,\rho_m\}$ such that $\ell_i(X) \in \beta_j^m$. Define now the integer $\ell_i'(X,H)$ by

$$\ell_i'(X,H) = \ell_{\ell_i(X)}(U_{\alpha_m}^\top HU_{\alpha_m}),$$

which, in fact, signifies the number of eigenvalues of $U_{\alpha_m}^\top HU_{\alpha_m}$ that are equal to $\lambda_{\ell_i(X)}(U_{\alpha_m}^\top HU_{\alpha_m})$, but are ranked before $\lambda_{\ell_i(X)}(U_{\alpha_m}^\top HU_{\alpha_m})$, including $\lambda_{\ell_i(X)}(U_{\alpha_m}^\top HU_{\alpha_m})$. As before, we often drop X and H from $\ell_i'(X,H)$ when the dependence of ℓ_i' on X and H can be seen clearly from the context. In summary, for any $i \in \{1,\ldots,n\}$, there are $m \in \{1,\ldots,r\}$ and $j \in \{1,\ldots,\rho_m\}$, for which we have, respectively,

$$i \in \alpha_m \quad \text{and} \quad \ell_i(X) \in \beta_j^m. \tag{15}$$

The following second-order estimate of eigenvalue functions was established in Torki [27, proposition 2.2] and has important consequences for second-order variational analysis of eigenvalue functions; see also Zhang et al. [28, proposition 2.1].

Proposition 2 (Second-Order Expansion of Eigenvalue Functions). *Assume that $X \in \mathbf{S}^n$ has the Eigenvalue Decomposition (4) for some $U \in \mathbf{O}^n(X)$ and that $H, W \in \mathbf{S}^n$. Let $\mu_1 > \cdots > \mu_r$ be distinct eigenvalues of X. Then, for any $t > 0$ sufficiently small and any $i \in \{1,\ldots,n\}$, we have*

$$\lambda_i(Y(t)) = \lambda_i(X) + t\lambda_{\ell_i}(U_{\alpha_m}^\top HU_{\alpha_m}) + \frac{1}{2}t^2\lambda_{\ell_i'}(R_{mj}{}^\top(U_{\alpha_m}^\top(W + 2H(\mu_m I - X)^\dagger H)U_{\alpha_m})R_{mj}) + o(t^2),$$

where $Y(t) := X + tH + \frac{1}{2}t^2W + o(t^2) \in \mathbf{S}^n$, and $R_{mj} := (Q_m)_{\beta_j^m}$ with Q_m and β_j^m taken from (13) and (15), respectively, and where the indices m and j come from (15).

In the framework of Proposition 2, we can conclude from (12) that for any $i \in \{1,\ldots,n\}$, the parabolic second-order directional derivative of the eigenvalue function $\lambda_i(\cdot)$ at X for H with respect to W, denoted $\lambda_i''(X;H,W)$, exists. Recall that the latter concept is defined by

$$\lambda_i''(X;H,W) = \lim_{t\downarrow 0}\frac{\lambda_i\left(X + tH + \frac{1}{2}t^2W\right) - \lambda_i(X) - t\lambda_i'(X;H)}{\frac{1}{2}t^2}.$$

According to Proposition 2, we can conclude further that

$$\lambda_i''(X;H,W) = \lambda_{\ell_i'}(R_{mj}{}^\top(U_{\alpha_m}^\top(W + 2H(\mu_m I - X)^\dagger H)U_{\alpha_m})R_{mj}). \tag{16}$$

Combining these with (12) brings us to the following estimate for the eigenvalue function $\lambda(\cdot)$ from (4), important for our development in this paper.

Corollary 1. *Assume that $X \in \mathbf{S}^n$ has the Eigenvalue Decomposition (4) for some $U \in \mathbf{O}^n(X)$ and that $H, W \in \mathbf{S}^n$. Then, for any $t > 0$ sufficiently small, we have*

$$\lambda\left(X + tH + \frac{1}{2}t^2W + o(t^2)\right) = \lambda(X) + t\lambda'(X;H) + \frac{1}{2}t^2\lambda''(X;H,W) + o(t^2). \tag{17}$$

We proceed with recalling some concepts utilized extensively in this paper. Given a nonempty set $C \subset \mathbf{X}$ with $\bar{x} \in C$, the tangent cone $T_C(\bar{x})$ to C at $\bar{x}$ is defined by

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

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 $\xi_+'(0) := \lim_{t\downarrow 0}[\xi(t) - \xi(0)]/t$ signifies the right derivative of ξ at 0. 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 $t\downarrow 0$ in the sense of the Painlevé-Kuratowski set convergence (cf. Rockafellar and Wets [25, definition 4.1]).

Given a function $f:\mathbf{X}\to\overline{\mathbf{R}}$, its domain is defined by $\operatorname{dom} f=\{x\in\mathbf{X}|f(x)<\infty\}$. The function f is called locally Lipschitz continuous around $\bar{x}$ *relative* to $C\subset\operatorname{dom} f$ with constant $\ell\geq 0$ if $\bar{x}\in C$ with $f(\bar{x})$ finite, 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 it is locally Lipschitz continuous around every $\bar{x}\in C$ 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. The subderivative function of f at $\bar{x}$, denoted by $\mathrm{d}f(\bar{x}):\mathbf{X}\to\overline{\mathbf{R}}$, is defined by

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

When f is convex, its subdifferential at $\bar{x}$ with $f(\bar{x})$ finite, denoted by $\partial f(\bar{x})$, is understood in the sense of convex analysis, namely, $v\in\partial f(\bar{x})$ if $f(x)\geq f(\bar{x})+\langle v,x-\bar{x}\rangle$ for any $x\in\mathbf{X}$. Given a nonempty convex set $C\subset\mathbf{X}$, its normal cones to C at $\bar{x}\in C$ is defined by $N_C(\bar{x})=\partial\delta_C(\bar{x})$. The second-order tangent set to $C\subset\mathbf{X}$ at $\bar{x}\in C$ for a tangent vector $w\in T_C(\bar{x})$ is given by

$$T_C^2(\bar{x},w)=\left\{u\in\mathbf{X}\,|\,\exists\, t_k\downarrow 0,\ u_k\to u \;\text{ as }\; k\to\infty \;\text{ with }\; \bar{x}+t_kw+\frac{1}{2}t_k^2u_k\in C\right\}. \tag{18}$$

A set C is called 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 arc $\xi:[0,\varepsilon]\to C$ with $\xi(0)=\bar{x}$, $\xi'_+(0)=w$, and $\xi''_+(0)=u$, where $\xi''_+(0):=\lim_{t\downarrow 0}[\xi(t)-\xi(0)-t\xi'_+(0)]/\frac{1}{2}t^2$. It is known that if C 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 $\mathbf{X}$ (cf. Bonnans and Shapiro [3, p. 163]). Below, we record a simple characterization of parabolic derivability of a set, used extensively in our paper.

Proposition 3. *Assume that $C\subset\mathbf{X}$, $\bar{x}\in C$, and $w\in T_C(\bar{x})$. Then, the following are equivalent*:

a. *C is parabolically derivable at $\bar{x}$ for w;*
b. *For any $u\in T_C^2(\bar{x},w)$, we find $\varepsilon>0$ such that*

$$\bar{x}+tw+\frac{1}{2}t^2u+o(t^2)\in C \quad \textit{for all} \;\; t\in[0,\varepsilon].$$

Proof. If (b) is satisfied, one can define $\xi(t)=\bar{x}+tw+\frac{1}{2}t^2u+o(t^2)$ for any $t\in[0,\varepsilon]$ with ε taken from (b). It is easy to see that $\xi(0)=\bar{x}$, $\xi'_+(0)=w$, and $\xi''_+(0)=u$, which confirm (a). Suppose that (a) holds and then pick $u\in T_C^2(\bar{x},w)$. Because C is parabolic drivable at $\bar{x}$ for w, we find $\varepsilon>0$ and an arc $\xi:[0,\varepsilon]\to C$ such that $\xi(0)=\bar{x}$, $\xi'_+(0)=w$, and $\xi''_+(0)=u$. Set $u(t)=(\xi(t)-\xi(0)-t\xi'_+(0))/\frac{1}{2}t^2$ for any $t\in[0,\varepsilon]$. It follows from $\xi''_+(0)=u$ that $u(t)\to u$ as $t\downarrow 0$. By the definition of $\xi''_+(0)$, we get

$$\xi(0)+t\xi'_+(0)+\frac{1}{2}t^2u(t)=\xi(t)\in C.$$

One other hand, one can express $\xi(t)$ equivalently as

$$\xi(t)=\xi(0)+t\xi'_+(0)+\frac{1}{2}t^2u+v(t) \quad \text{with} \;\; v(t):=\frac{1}{2}t^2(u(t)-u).$$

Clearly, we have $v(t)=o(t^2)$, which proves (b) and, hence, completes the proof. □

3. Subderivatives of Spectral Functions

In this section, we present two important results about the spectral functions, central to our developments in this paper. The first one presents a counterpart of the estimate in (3) for symmetric functions in Proposition 4. The second one presents a chain rule for the subderivative of spectral functions in Theorem 1. To state the former about symmetric functions, recall that a function $\theta:\mathbf{R}^n\to\overline{\mathbf{R}}$ is called symmetric if for every $x\in\mathbf{R}^n$ and every $n\times n$ permutation matrix Q, we have $\theta(Qx)=\theta(x)$. Recall also that Q is a permutation matrix if all its components are either 0 or 1 and each row and each column has exactly one nonzero element. We denote by $\mathbf{P}^n$ the set of all $n\times n$ permutation matrices. As pointed out before, for any spectral function $g:\mathbf{S}^n\to\overline{\mathbf{R}}$, there exists a symmetric function

$\theta : \mathbf{R}^n \to \overline{\mathbf{R}}$ satisfying (1). Indeed, θ can be chosen as the restriction of g to diagonal matrices, namely,

$$\theta(x) = g(\mathrm{Diag}\,(x)) \quad \text{for all } \; x \in \mathbf{R}^n. \tag{19}$$

A set $C \subset \mathbf{S}^n$ is called a spectral set if δ_C is a spectral function. Likewise, $\Theta \subset \mathbf{R}^n$ is called a symmetric set if δ_Θ is a symmetric function. Similar to (1), it is easy to see that for any spectral set $C \subset \mathbf{S}^n$, there exists a symmetric set $\Theta \subset \mathbf{R}^n$ such that

$$C = \{X \in \mathbf{S}^n \,|\, \lambda(X) \in \Theta\}, \tag{20}$$

where Θ can be chosen as

$$\Theta = \{x \in \mathbf{R}^n \,|\, \mathrm{Diag}\,(x) \in C\}. \tag{21}$$

The Composite Forms (1) and (19) readily imply, respectively, that

$$\mathrm{dom}\; g = \{X \in \mathbf{S}^n \,|\, \lambda(X) \in \mathrm{dom}\; \theta\} \quad \text{and} \quad \mathrm{dom}\; \theta = \{x \in \mathbf{R}^n \,|\, \mathrm{Diag}\,(x) \in \mathrm{dom}\; g\}. \tag{22}$$

Next, we are going to justify a similar estimate as (3) for domains of symmetric functions, which allows us to show via the established theory for composite functions in Mohammadi and Sarabi [20] that second-order variational properties of spectral functions are inherited by symmetric functions.

Proposition 4. *Let $g : \mathbf{S}^n \to \overline{\mathbf{R}}$ be a spectral function, represented by (1). Then, for any $x \in \mathbf{R}^n$, we have*

$$\mathrm{dist}(x,\ \mathrm{dom}\; \theta) = \mathrm{dist}(\mathrm{Diag}\,(x),\ \mathrm{dom}\; g), \tag{23}$$

where θ is taken by (1).

Proof. For any $x \in \mathbf{R}^n$, we know that there exist a permutation matrix $P \in \mathbf{O}^n$ such that $\lambda(\mathrm{Diag}\,(x)) = Px$. Because θ is a symmetric function, dom θ is a symmetric set. Thus, for any $X \in \mathrm{dom}\; g$, we have $P^\top \lambda(X) \in \mathrm{dom}\; \theta$. This, coupled with (5), leads us to

$$\begin{aligned} \|\mathrm{Diag}\,(x) - X\| \geq \|\lambda(\mathrm{Diag}\,(x)) - \lambda(X)\| &= \|Px - \lambda(X)\| \\ &= \|x - P^\top \lambda(X)\| \geq \mathrm{dist}(x,\ \mathrm{dom}\; \theta), \end{aligned}$$

for all $X \in \mathrm{dom}\; g$, which, in turn, brings us to

$$\mathrm{dist}(x,\ \mathrm{dom}\; \theta) \leq \mathrm{dist}(\mathrm{Diag}\,(x),\ \mathrm{dom}\; g).$$

To prove the opposite inequality, pick any $y \in \mathrm{dom}\; \theta$. By (22), we get $\mathrm{Diag}\,(y) \in \mathrm{dom}\; g$, which implies that

$$\|x - y\| = \|\mathrm{Diag}\,(x) - \mathrm{Diag}\,(y)\| \geq \mathrm{dist}(\mathrm{Diag}\,(x),\ \mathrm{dom}\; g).$$

Combining these clearly justifies (23). □

Note that the identity in (3) allows us to show that second-order variational properties of a symmetric function θ from (1) are disseminated to the spectral function g. Appealing to (23), we will show in the coming sections that those variational properties of the spectral function g are inherited by the symmetric function θ from (1). This will be achieved by using the second-order variational theory in Mohammadi and Sarabi [20] for the Composite Form (19). Note also that the results in Mohammadi and Sarabi [20] were proven under a constraint qualification, which is similar to (2). According to (23), such a constraint qualification automatically holds for the Composite Form (19). Moreover, the inner mapping $x \longmapsto \mathrm{Diag}\,(x)$ in this composite form is twice continuously differentiable, which allows us to exploit the results in Mohammadi and Sarabi [20] and Mohammadi et al. [22].

Proposition 5. *Let $g : \mathbf{S}^n \to \overline{\mathbf{R}}$ be a spectral function, represented by (1), and let the symmetric function θ, taken from (1), be locally Lipschitz continuous relative to its domain. Then, for any $X \in \mathbf{S}^n$ with $g(X)$ finite and any $v \in \mathbf{R}^n$, we have*

$$\mathrm{d}\theta(\lambda(X))(v) = \mathrm{d}g(\mathrm{Diag}\,(\lambda(X)))(\mathrm{Diag}\,(v)). \tag{24}$$

In particular, if $g = \delta_C$, where $C \subset \mathbf{S}^n$ is a spectral set, then we get for any $X \in C$ that

$$T_\Theta(\lambda(X)) = \{v \in \mathbf{R}^n \,|\, \mathrm{Diag}\,(v) \in T_C(\mathrm{Diag}\,(\lambda(X))\}, \tag{25}$$

where Θ is taken from (20).

Proof. It follows from (1) that the symmetric function θ satisfies (19), which means that θ can be represented as a composite function of g and the linear mapping $x \longmapsto \mathrm{Diag}\,(x)$ with $x \in \mathbf{R}^n$. We also deduce from the imposed assumption on θ and the inequality in (5) that g is locally Lipschitz continuous relative to its domain. This, together with (23) and Mohammadi et al. [22, theorem 3.4], justifies (24). To justify (25), recall the representation Θ from (21), which can be equivalently expressed as $\delta_\Theta(x) = \delta_C(\mathrm{Diag}\,(x))$ for any $x \in \mathbf{R}^n$. The claim equality in (25) results from (24) and the fact that $\mathrm{d}\theta(\lambda(X)) = \delta_{T_\Theta(\lambda(X))}$ and $\mathrm{d}g(\mathrm{Diag}\,(\lambda(X))) = \delta_{T_C(\mathrm{Diag}\,(\lambda(X)))}$. □

Remark 1 (Symmetric Property of Subderivatives). If $\theta:\mathbf{R}^n\to\overline{\mathbf{R}}$ is a symmetric function and $X\in\mathbf{S}^n$ with $\theta(\lambda(X))$ finite, one may wonder whether the subderivative $\mathrm{d}\theta(\lambda(X))$ is a symmetric function. This can be easily disproven by taking $\theta=\delta_{\mathbf{R}^n_-}$ and $X\in\mathbf{S}^n$ such that all its eigenvalues are not the same; see Example 1 for more details. Although this may seem disappointing, we can show that $\mathrm{d}\theta(\lambda(X))$ is a symmetric function with respect to a subset of $\mathbf{P}^n$. Indeed, assume that $\mathbf{P}^n_X$ is the set of all $n\times n$ block diagonal matrices in the form $Q=\mathrm{Diag}\,(P_1,\ldots,P_r)$, where $P_m\in\mathbf{R}^{|\alpha_m|\times|\alpha_m|}$ is a permutation matrix for any $m=1,\ldots,r$ with α_m taken from (7) and r being the number of distinct eigenvalues of X. It is clear that $\mathbf{P}^n_X\subset\mathbf{P}^n$ and that if $Q\in\mathbf{P}^n_X$, then we have $Q\lambda(X)=\lambda(X)$. Moreover, for any $v\in\mathbf{R}^n$ and $Q\in\mathbf{P}^n_X$, we get

$$\begin{aligned}\mathrm{d}\theta(\lambda(X))(v)&=\liminf_{\substack{t\downarrow 0\\ v'\to v}}\frac{\theta(\lambda(X)+tv')-\theta(\lambda(X))}{t}\\&=\liminf_{\substack{t\downarrow 0\\ v'\to v}}\frac{\theta(\lambda(X)+tQv')-\theta(\lambda(X))}{t}\\&\geq\liminf_{\substack{t\downarrow 0\\ w\to Qv}}\frac{\theta(\lambda(X)+tw)-\theta(\lambda(X))}{t}=\mathrm{d}\theta(\lambda(X))(Qv).\end{aligned}$$

Because $Q^{-1}=\mathrm{Diag}\,(P_1^{-1},\ldots,P_r^{-1})$, we can show similarly that $\mathrm{d}\theta(\lambda(X))(v)\leq\mathrm{d}\theta(\lambda(X))(Qv)$ for any $v\in\mathbf{R}^n$ and $Q\in\mathbf{P}^n_X$, which leads us to

$$\mathrm{d}\theta(\lambda(X))(v)=\mathrm{d}\theta(\lambda(X))(Qv)\quad\text{for all }\; v\in\mathbf{R}^n,\ Q\in\mathbf{P}^n_X,$$

demonstrating that $\mathrm{d}\theta(\lambda(X))$ is a symmetric function with respect to $\mathbf{P}^n_X$.

We proceed by proving a chain rule for subderivatives of spectral functions. We begin with recalling a useful characterization of the subdifferential of the spectral functions.

Proposition 6. *Assume that $\theta:\mathbf{R}^n\to\overline{\mathbf{R}}$ is a proper, lower semicontinuous (lsc), convex, and symmetric function. Then, the following properties are equivalent:*

a. $Y\in\partial(\theta\circ\lambda)(X)$;

b. $\lambda(Y)\in\partial\theta(\lambda(X))$ *and the matrices X and Y have simultaneous ordered spectral decomposition, meaning that there exists* $U\in\mathbf{O}^n(X)\cap\mathbf{O}^n(Y)$ *such that*

$$X=U\Lambda(X)U^\top\quad\textit{and}\quad Y=U\Lambda(Y)U^\top,$$

where $\Lambda(X)=\mathrm{Diag}\,(\lambda(X))$ *and* $\Lambda(Y)=\mathrm{Diag}\,(\lambda(Y))$.

Proof. It follows from Borwein and Lewis [4, corollary 5.2.3] that $\theta\circ\lambda$ is lsc and convex if and only if θ is lsc and convex. The claimed equivalence then results from Borwein and Lewis [4, theorem 5.2.4]. □

Given a matrix $X\in\mathbf{S}^n$ with r distinct eigenvalues and the index sets α_m, $m=1,\ldots,r$, from (7), recall that $\cup_{m=1}^r\alpha_m=\{1,\ldots,n\}$. In what follows, we partition a vector $p\in\mathbf{R}^n$ into $(p_{\alpha_1},\ldots,p_{\alpha_r})$, where $p_{\alpha_m}\in\mathbf{R}^{|\alpha_m|}$ for any $m=1,\ldots,r$.

Theorem 1 (Subderivatives of Spectral Functions). *Let $\theta:\mathbf{R}^n\to\overline{\mathbf{R}}$ be a symmetric function and let $X\in\mathbf{S}^n$ with $(\theta\circ\lambda)(X)$ finite. If θ is either lsc and convex with $\partial\theta(\lambda(X))\neq\emptyset$ or locally Lipschitz continuous around $\lambda(X)$ relative to its domain, then for all $H\in\mathbf{S}^n$, we have*

$$\mathrm{d}(\theta\circ\lambda)(X)(H)=\mathrm{d}\theta(\lambda(X))(\lambda'(X;H)).\tag{26}$$

Proof. Pick any $H\in\mathbf{S}^n$ and deduce from Proposition 1 that $\lambda'(X;.)$ is a Lipschitz-continuous and positively homogeneous function. Moreover, we have $\lambda'(X;E)+O(t^2\|E\|^2)/t\to\lambda'(X;H)$ as $t\downarrow 0$ and $E\to H$. This and the definition of subderivative give us the relationships

$$\begin{aligned}\mathrm{d}(\theta\circ\lambda)(X)(H)&=\liminf_{\substack{t\downarrow 0\\ E\to H}}\frac{\theta(\lambda(X+tE))-\theta(\lambda(X))}{t}\\&=\liminf_{\substack{t\downarrow 0\\ E\to H}}\frac{\theta(\lambda(X)+t\lambda'(X;E)+O(t^2\|E\|^2))-\theta(\lambda(X))}{t}\\&=\liminf_{\substack{t\downarrow 0\\ E\to H}}\frac{\theta(\lambda(X)+t[\lambda'(X;E)+O(t^2\|E\|^2)/t])-\theta(\lambda(X))}{t}\\&\geq\mathrm{d}\theta(\lambda(X))(\lambda'(X;H)),\end{aligned}\tag{27}$$

which verify the inequality "$\geq$" in (26). To justify the opposite inequality in (27), observe that if $d\theta(\lambda(X))(\lambda'(X;H)) = \infty$, the latter inequality clearly holds. Thus, assume that $d\theta(\lambda(X))(\lambda'(X;H)) < \infty$. If θ is lsc and convex, $\theta \circ \lambda$ is lsc and convex due to Borwein and Lewis [4, corollary 5.2.3]. Moreover, it follows from Lewis [18, theorem 6] and $\partial\theta(\lambda(X)) \neq \emptyset$ that $\partial(\theta \circ \lambda)(X) \neq \emptyset$. Thus, it follows from Rockafellar and Wets [25, theorem 8.30] that $d(\theta \circ \lambda)(X)(H) = \sup_{Y \in \partial(\theta\circ\lambda)(X)} \langle Y, H\rangle$. Let $\varepsilon > 0$ and choose $Y \in \partial(\theta \circ \lambda)(X)$ such that

$$d(\theta \circ \lambda)(X)(H) \leq \varepsilon + \langle Y, H\rangle.$$

Because $Y \in \partial(\theta \circ \lambda)(X)$, it follows from Proposition 6 that there is $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$ such that $\lambda(Y) \in \partial\theta(\lambda(X))$. Set $\Lambda(Y) := U^\top Y U$ and use Fan's inequality to conclude

$$\begin{aligned} d(\theta \circ \lambda)(X)(H) \leq \varepsilon + \langle Y, H\rangle = \varepsilon + \langle \Lambda(Y), U^T H U\rangle &= \varepsilon + \sum_{m=1}^{r} \langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H U_{\alpha_m}\rangle \\ &\leq \varepsilon + \sum_{m=1}^{r} \langle \lambda(Y)_{\alpha_m}, \lambda(U_{\alpha_m}^\top H U_{\alpha_m})\rangle \\ &= \varepsilon + \langle \lambda(Y), \lambda'(X;H)\rangle \\ &\leq \varepsilon + d\theta(\lambda(X))(\lambda'(X;H)), \end{aligned}$$

where the last inequality results from the fact that θ is convex and $\lambda(Y) \in \partial\theta(\lambda(X))$. Letting $\varepsilon \downarrow 0$, we get the opposite inequality in (27), which proves (26) in this case. Suppose now that θ is locally Lipschitz continuous around $\lambda(X)$ relative to its domain. To prove the opposite inequality in (27), by definition, there exist sequences $t_k \downarrow 0$ and $v_k \to \lambda'(X;H)$ such that

$$d\theta(\lambda(X))(\lambda'(X;H)) = \lim_{k\to\infty} \frac{\theta(\lambda(X) + t_k v_k) - \theta(\lambda(X))}{t_k}. \tag{28}$$

Because $d\theta(\lambda(X))(\lambda'(X;H)) < \infty$, we can assume without loss of generality that $\lambda(X) + t_k v_k \in \text{dom}\,\theta$ for all $k \in \mathbf{N}$. Take the function g from (1) and appeal to (3) to get

$$\text{dist}(X + t_k H,\ \text{dom}\, g) = \text{dist}(\lambda(X + t_k H),\ \text{dom}\,\theta), \quad k \in \mathbf{N},$$

which, in turn, brings us to the relationships

$$\begin{aligned} \text{dist}\left(H, \frac{\text{dom}\, g - X}{t_k}\right) &= \frac{1}{t_k}\ \text{dist}(\lambda(X) + t_k\lambda'(X,H) + O(t_k^2),\ \text{dom}\,\theta) \\ &\leq \frac{1}{t_k}\ \|\lambda(X) + t_k\lambda'(X,H) + O(t_k^2) - \lambda(X) - t_k v_k\| \\ &= \left\|\lambda'(X;H) - v_k + \frac{O(t_k^2)}{t_k}\right\| \ \text{for all}\ \ k \in \mathbf{N}. \end{aligned}$$

So, for each $k \in \mathbf{N}$, we find a matrix $E_k \in \mathbf{S}^n$ such that $X + t_k E_k \in \text{dom}\, g$ and

$$\|H - E_k\| < \left\|\lambda'(X;H) - v_k + \frac{O(t_k^2)}{t_k}\right\| + \frac{1}{k},$$

which, in turn, yields $E_k \to H$ as $k \to \infty$. Combining these with (28) and (11), we arrive at

$$\begin{aligned} d\theta(\lambda(X))(\lambda'(X;H)) &= \lim_{k\to\infty}\left[\frac{g(X + t_k E_k) - g(X)}{t_k} + \frac{\theta(\lambda(X) + t_k v_k) - \theta(\lambda(X + t_k E_k))}{t_k}\right] \\ &\geq \liminf_{k\to\infty} \frac{g(X + t_k E_k) - g(X)}{t_k} - \kappa \lim_{k\to\infty}\left\|\frac{\lambda(X + t_k E_k) - \lambda(X)}{t_k} - v_k\right\| \\ &\geq dg(X)(H) - \kappa \lim_{k\to\infty}\left\|\lambda'(X;E_k) + \frac{O(t_k^2)}{t_k} - v_k\right\| = dg(X)(H), \end{aligned}$$

where $\kappa \geq 0$ is a Lipschitz constant of θ around $\lambda(X)$ relative to its domain. This verifies the inequality "$\leq$" in (26) and completes the proof of the theorem. □

As an immediate conclusion of Theorem 1, we obtain a simple representation of tangent cones to spectral sets.

Corollary 2 (Tangent Cone to the Spectral Sets). *Let C be a spectral set represented by* (20). *Then, for any $X \in C$, we have*

$$T_C(X) = \{H \in \mathbf{S}^n \,|\, \lambda'(X;H) \in T_\Theta(\lambda(X))\}.$$

Proof. Taking the symmetric set Θ from (20), we can apply Theorem 1 for the symmetric function δ_Θ. The claimed representation of the tangent cone to C at X follows from the facts that $\mathrm{d}\delta_\Theta(X) = \delta_{T_C(X)}$ and $\mathrm{d}\delta_\Theta(\lambda(X)) = \delta_{T_\Theta(\lambda(X))}$. □

Example 1 (Tangent Cone to $\mathbf{S}^n_-$). Suppose that $\mathbf{S}^n_-$ stands for the cone of all $n \times n$ symmetric and negative semidefinite matrices. This cone is a spectral set, and

$$\mathbf{S}^n_- = \{X \in \mathbf{S}^n \,|\, \lambda(X) \in \mathbf{R}^n_-\}.$$

Take $X \in \mathbf{S}^n_-$ and assume that $\mu_1 > \cdots > \mu_r$ are its distinct eigenvalues. If $\mu_1 < 0$, then we clearly have $\lambda(X) \in \operatorname{int} \mathbf{R}^n_-$ and, hence, $T_{\mathbf{R}^n_-}(\lambda(X)) = \mathbf{R}^n$. Using this, together with Corollary 2, we get $T_{\mathbf{S}^n_-}(X) = \mathbf{S}^n$. If $\mu_1 = 0$, then we obtain $T_{\mathbf{R}^n_-}(\lambda(X)) = \mathbf{R}^{|\alpha_1|}_- \times \mathbf{R}^{n-|\alpha_1|}$, where α_1 is defined by (7). Appealing to Corollary 2 tells us that

$$T_{\mathbf{S}^n_-}(X) = \{H \in \mathbf{S}^n \,|\, \lambda_1(U^\top_{\alpha_1} H U_{\alpha_1}) \le 0\}, \tag{29}$$

where U is taken from (4).

The tangent cone description for the set $\mathbf{S}^n_-$ in (29) can be alternatively obtained by Bonnans and Shapiro [3, proposition 2.61]. Indeed, it is easy to see that $\mathbf{S}^n_- = \{X \in \mathbf{S}^n \,|\, \lambda_1(X) \le 0\}$, in which λ_1 is known to be a convex function (cf. Rockafellar and Wets [25, exercise 2.54]). Obviously, we can find $X \in \mathbf{S}^n$ with $\lambda_1(X) < 0$, a condition known as the *Slater* condition and assumed in Bonnans and Shapiro [3, proposition 2.61]. In contrast, our approach relies upon the metric subregularity, automatically satisfied for spectral sets. This allows us to calculate the tangent cone of spectral sets, even if the Slater condition fails therein, as the following example demonstrates.

Example 2 (Failure of the Slater Condition in Spectral Sets). Assume that $k \in \mathbf{N}$ and consider the set

$$C = \left\{ X \in \mathbf{S}^n_+ \,\middle|\, \sum_{i=1}^n \lambda_i^k(X) = 1 \right\}. \tag{30}$$

This is clearly a spectral subset of $\mathbf{S}^n$. When $k = 1$, the set C is called spectahedron. Note that C can be represented in the form of (20) with the symmetric set Θ defined by

$$\Theta = \left\{ (z_1, \ldots, z_n) \in \mathbf{R}^n \,\middle|\, \sum_{i=1}^n z_i^k = 1,\, z_i \ge 0 \text{ for all } \ i \in \{1, \ldots, n\} \right\}. \tag{31}$$

Set $\Phi(z) = (\sum_{i=1}^n z_i^k - 1, z_1, \ldots, z_n)$ with $z = (z_1, \ldots, z_n)$ and $D = \{0\} \times \mathbf{R}^n_+$ and observe that

$$\Theta = \{z \in \mathbf{R}^n \,|\, \Phi(z) \in D\}. \tag{32}$$

We claim now that

$$N_D(\Phi(z)) \cap \ker \nabla\Phi(z)^* = \{0\}, \tag{33}$$

for any $z \in \Theta$. To justify it, pick $z = (z_1, \ldots, z_n) \in \Theta$ and assume that $(b_0, \ldots, b_n) \in N_D(\Phi(z)) \cap \ker \nabla\Phi(z)^*$. This implies that $b_i z_i = 0$ and $k z_i^{k-1} b_0 + b_i = 0$ for any $i = 1, \ldots, n$. It is not hard to see that these conditions lead us to $b_i = 0$ for any $i = 0, \ldots, n$, which proves our claim. Take $X \in C$ and define the active index set $I(\lambda(X)) = \{i \in \{1, \ldots, n\} \,|\, \lambda_i(X) = 0\}$. It follows from Rockafellar and Wets [25, theorem 6.14] and (33) that the tangent cone to Θ at $\lambda(X)$ can be calculated as

$$T_\Theta(\lambda(X)) = \left\{ (w_1, \ldots, w_n) \in \mathbf{R}^n \,\middle|\, \sum_{i=1}^n w_i \lambda_i^{k-1}(X) = 0,\, w_i \ge 0 \text{ for all } i \in I(\lambda(X)) \right\}.$$

Appealing to Corollary 2 tells us that

$$T_C(X) = \left\{ H \in \mathbf{S}^n \,\middle|\, \sum_{i=1}^n \lambda_i'(X;H) \lambda_i^{k-1}(X) = 0,\ \ \lambda_i'(X;H) \ge 0 \text{ for all } i \in I(\lambda(X)) \right\}. \tag{34}$$

Note that, due to the presence of the equality constraint, the Slater condition fails for C, and, thus, Bonnans and Shapiro [3, proposition 2.61] can't be applied.

4. Parabolic Epi-Differentiability of Spectral Functions

The main objective of this section is to provide a systematic study of two important second-order variational properties of spectral sets and functions: (1) parabolic derivability; and (2) a chain rule for parabolic subderivatives. To achieve these goals, we begin with justifying that certain second-order approximations of spectral sets enjoy an outer Lipschitzian property, which is central to our developments in this section. Suppose that $C \subset \mathbf{S}^n$ is a spectral set with the representation in (20) and that $X \in C$ and $H \in T_C(X)$. Define the set-valued mapping $S_H : \mathbf{R}^n \rightrightarrows \mathbf{S}^n$ via the second-order tangent set to the symmetric set Θ in (20) by

$$S_H(p) := \{W \in \mathbf{S}^n \,|\, \lambda''(X; H, W) + p \in T^2_\Theta(\lambda(X), \lambda'(X; H))\}. \tag{35}$$

For any parameter $p \in \mathbf{R}^n$, the set-valued mapping $S_H(p)$ presents a second-order tangential approximation of the Spectral Set (20) at X for H. Note that by Corollary 2, the condition $H \in T_C(X)$ amounts to $\lambda'(X; H) \in T_\Theta(\lambda(X))$, which is required in the definition of the second-order tangent set to Θ at $\lambda(X)$ in (35); see (18). Note also that reducing the Estimate (3), which was stated for the spectral function in (1), to the spectral set C in (20) gives us the estimate

$$\mathrm{dist}(X, C) = \mathrm{dist}(\lambda(X), \Theta) \quad \text{for all} \ \ X \in \mathbf{S}^n, \tag{36}$$

which will be utilized broadly in this section.

Proposition 7 (Uniform Outer Lipschitzian Property of S_H). *Assume that $C \subset \mathbf{S}^n$ is a spectral set with the Representation (20) and that $X \in C$ and $H \in T_C(X)$. Then, the mapping S_H in (35) enjoys the following uniform outer Lipschitzian property at the origin*:

$$S_H(p) \subset S_H(0) + \|p\|\mathbf{B} \ \textit{for all} \ \ p \in \mathbf{R}^n. \tag{37}$$

Proof. Let $p \in \mathbf{R}^n$ and pick then $W \in S_H(p)$. It follows from (35) that $\lambda''(X; H, W) + p \in T^2_\Theta(\lambda(X), \lambda'(X; H))$. By (18), there exists a sequence $t_k \downarrow 0$ such that

$$\lambda(X) + t_k\lambda'(X; H) + \frac{1}{2}t_k^2\lambda''(X; H, W) + \frac{1}{2}t_k^2 p + o(t_k^2) \in \Theta \quad \text{for all} \ \ k \in \mathbf{N}.$$

For any k sufficiently large, we conclude from (17) that

$$\lambda\left(X + t_kH + \frac{1}{2}t_k^2W\right) = \lambda(X) + t_k\lambda'(X; H) + \frac{1}{2}t_k^2\lambda''(X; H, W) + o(t_k^2),$$

which, in turn, implies via (36) that

$$\mathrm{dist}\left(X + t_kH + \frac{1}{2}t_k^2W, C\right) = \mathrm{dist}\left(\lambda\left(X + t_kH + \frac{1}{2}t_k^2W\right), \Theta\right) \le \frac{1}{2}t_k^2\|p\| + o(t_k^2).$$

This ensures the existence of a matrix $Y_k \in C$ such that

$$\|D_k\| \le \frac{1}{2}\left(\|p\| + \frac{o(t_k^2)}{t_k^2}\right) \ \text{with} \ \ D_k := \frac{X + t_kH + \frac{1}{2}t_k^2W - Y_k}{t_k^2}.$$

Passing to a subsequence, if necessary, ensures the existence of $D \in \mathbf{S}^n$ such that $D_k \to D$ as $k \to \infty$. This yields the estimate

$$\|D\| \le \frac{1}{2}\|p\|. \tag{38}$$

It follows from $X + t_kH + \frac{1}{2}t_k^2W - t_k^2D_k = Y_k \in C$ and (20) that $\lambda\left(X + t_kH + \frac{1}{2}t_k^2W - t_k^2D_k\right) \in \Theta$. Taking into account (17), we get for any $k \in \mathbf{N}$ sufficiently large that

$$\lambda\left(X + t_kH + \frac{1}{2}t_k^2W - t_k^2D_k\right) = \lambda(X) + t_k\lambda'(X; H) + \frac{1}{2}t_k^2\lambda''(X; H, W - 2D) + o(t_k^2) \in \Theta.$$

By the definition of the second-order tangent set, we arrive at

$$\lambda''(X; H, W - 2D) \in T^2_\Theta(\lambda(X), \lambda(X; H)),$$

which yields $W - 2D \in S_H(0)$. This, combined with (38), justifies the claimed inclusion in (37) and thus completes the proof. $\square$

The outer Lipschitzian property for second-order tangential approximations appeared first in Mohammadi et al. [21, theorem 4.3] for sets C as the one in (20) with the eigenvalue function $\lambda(\cdot)$ replaced with a twice differentiable function, under an adaptation of the Constraint Qualification (2) for this setting. Proposition 7 demonstrates that the latter result can be achieved without the assumed twice differentiability in Mohammadi et al. [21] when we still have a second-order expansion for functions in our settings.

Next, we are going to achieve a chain rule for second-order tangent sets of spectral sets, which heavily relies upon Proposition 7. First, we recall Mohammadi et al. [21, theorem 4.5], where a similar result was proven for constraint systems in finite dimensional Hilbert spaces.

Proposition 8. *Let D be a closed subset of $\mathbf{Y}$ and let $\Omega = \{x \in \mathbf{X} | \Phi(x) \in D\}$, where $\Phi : \mathbf{X} \to \mathbf{Y}$ is a twice differentiable function between two Euclidean spaces, and $\bar{x} \in \Omega$. Suppose further that there are $\kappa \geq 0$ and $\varepsilon > 0$ such that the estimate*

$$\mathrm{dist}(x, \Omega) \leq \kappa\, \mathrm{dist}(\Phi(x), D) \ \text{ for all } \ x \in \mathbf{B}_\varepsilon(\bar{x}), \tag{39}$$

holds. Then, for all $w \in T_\Omega(\bar{x})$, we have

$$T^2_\Omega(\bar{x}, w) = \{u \in \mathbf{X} | \nabla\Phi(\bar{x})u + \nabla^2\Phi(\bar{x})(w, w) \in T^2_D(\Phi(\bar{x}), \nabla\Phi(\bar{x})w)\}. \tag{40}$$

If, furthermore, the set D is parabolically derivable at $\Phi(\bar{x})$ for $\nabla\Phi(\bar{x})w$, then the constraint set Ω is parabolically derivable at $\bar{x}$ for w.

Proof. The equality in (40) was justified in Mohammadi et al. [21, theorem 4.5] under an extra assumption that the set D is regular in the sense of Rockafellar and Wets [25, definition 6.4]; see Gfrerer et al. [15, proposition 5] for an extension of Mohammadi et al. [21, theorem 4.5] without the regularity assumption on D. Note that the directional metric subregularity used in Gfrerer et al. [15, proposition 5] is weaker than (39) in general. However, it was shown in Gfrerer and Outrata [14, lemma 2.8(ii)] that metric subregularity at any direction is equivalent to (39).

To prove parabolic derivability of Ω at $\bar{x}$ for w, pick $u \in T^2_\Omega(\bar{x}, w)$. It follows from the proof of Gfrerer et al. [15, proposition 5] that there is a positive constant ℓ such that

$$\mathrm{dist}\left(u, \frac{\Omega - \bar{x} - tw}{\frac{1}{2}t^2}\right) \leq \ell\, \mathrm{dist}\left(\nabla\Phi(\bar{x})u + \nabla^2\Phi(\bar{x})(w, w), \frac{D - \Phi(\bar{x}) - t\nabla\Phi(\bar{x})w}{\frac{1}{2}t^2}\right) + \frac{o(t^2)}{t^2}, \tag{41}$$

for any t sufficiently small that $t \downarrow 0$. Because $u \in T^2_\Omega(\bar{x}, w)$, we conclude from (40) that $\nabla\Phi(\bar{x})u + \nabla^2\Phi(\bar{x})(w, w) \in T^2_D(\Phi(\bar{x}), \nabla\Phi(\bar{x})w)$, which together with parabolic derivability of D at $\Phi(\bar{x})$ for $\nabla\Phi(\bar{x})w$ implies via Rockafellar and Wets [25, corollary 4.7] that

$$\mathrm{dist}\left(\nabla\Phi(\bar{x})u + \nabla^2\Phi(\bar{x})(w, w), \frac{D - \Phi(\bar{x}) - t\nabla\Phi(\bar{x})w}{\frac{1}{2}t^2}\right) \to 0 \quad \text{and} \quad t \downarrow 0.$$

This, coupled with (41), confirms that for any sufficiently small t, there exists $u(t) \in (\Omega - \bar{x} - tw)/\frac{1}{2}t^2$ such that $u(t) \to u$ as $t \downarrow 0$. Define the arc $\xi(t) := -\bar{x} + tw + \frac{1}{2}t^2 u(t)$ and observe that $\xi(0) = \bar{x}$, $\xi'_+(0) = w$, and $\xi''_+(0) = u$. To finish the proof, we need to show that $T^2_\Omega(\bar{x}, w) \neq \emptyset$, which was already established in Gfrerer et al. [15, corollary 1] under the metric subregularity condition in (39). Combining these confirms that Ω is parabolically derivable at $\bar{x}$ for w and, hence, completes the proof. □

Theorem 2 (Second-Order Tangent Sets of Spectral Sets). *Assume that $C \subset \mathbf{S}^n$ is a spectral set with the representation in (20) and that $X \in C$ and $H \in T_C(X)$. Then, we have*

$$T^2_C(X, H) = \{W \in \mathbf{S}^n | \lambda''(X; H, W) \in T^2_\Theta(\lambda(X), \lambda'(X; H))\}, \tag{42}$$

where Θ is taken from (20). Moreover, the following properties are satisfied.

a. *If the symmetric set Θ is parabolically derivable at $\lambda(X)$ for $\lambda'(X; H)$, then C is parabolically derivable at X for H.*

b. *If the symmetric set C is parabolically derivable at X for any $H \in T_C(X)$, then Θ is parabolically derivable at $\lambda(X)$ for any $v \in T_\Theta(\lambda(X))$.*

Proof. First note from Corollary 2 that the condition $H \in T_C(X)$ amounts to $\lambda'(X; H) \in T_\Theta(\lambda(X))$. Let $W \in \mathbf{S}^n$. Employing (36) and (17) tells us that for any $t > 0$ sufficiently small, we have

$$\begin{aligned}\mathrm{dist}\left(X + tH + \frac{1}{2}t^2 W, C\right) &= \mathrm{dist}\left(\lambda\left(X + tH + \frac{1}{2}t^2 W\right), \Theta\right) \\ &= \mathrm{dist}\left(\lambda(X) + t\lambda'(X; H) + \frac{1}{2}t^2 \lambda''(X; H, W) + o(t^2), \Theta\right).\end{aligned} \tag{43}$$

Take $W \in T^2_C(X,H)$. By (18), there exists a sequence $t_k \downarrow 0$ such that $X + t_kH + \frac{1}{2}t_k^2W + o(t_k^2) \in C$. By (43), we get $\lambda(X) + t_k\lambda'(X;H) + \frac{1}{2}t_k^2\lambda''(X;H,W) + o(t_k^2) \in \Theta$, which clearly demonstrates that $\lambda''(X;H,W) \in T^2_\Theta(\lambda(X), \lambda'(X;H))$ and thus proves the inclusion "$\subset$" in (42). The opposite inclusion in (42) can be established via a similar argument and (43), which proves the claimed representation of the second-order tangent set to C in (42).

To prove (a), suppose that the symmetric set Θ is parabolically derivable at $\lambda(X)$ for $\lambda'(X;H)$. To justify the same property for C at X for H, pick $W \in T^2_C(X,H)$. By (42), we obtain $\lambda''(X;H,W) \in T^2_\Theta(\lambda(X), \lambda'(X;H))$. Because Θ is parabolically derivable at $\lambda(X)$ for $\lambda'(X;H)$, we conclude from Proposition 3 that there exists $\varepsilon > 0$ such that for all $t \in [0,\varepsilon]$, we have

$$\lambda(X) + t\lambda'(X;H) + \frac{1}{2}t^2\lambda''(X;H,W) + o(t^2) \in \Theta.$$

Reducing $\varepsilon > 0$ if necessary, pick $t \in [0,\varepsilon]$ and conclude from (43) that $X + tH + \frac{1}{2}t^2W + o(t^2) \in C$. Defining the arc $\xi : [0,\varepsilon] \to C$ by $\xi(t) = X + tH + \frac{1}{2}t^2W + o(t^2)$ for $t \in [0,\varepsilon]$, we can readily see that $\xi(0) = X$, $\xi'_+(0) = H$, and $\xi''_+(0) = W$. To finish the proof of parabolic derivability of C at X for H, it remains to show that $T^2_C(X,H) \neq \emptyset$. To this end, pick $Z \in \mathbf{S}^n$ and $y \in T^2_\Theta(\lambda(X), \lambda'(X;H))$. In fact, such y exists because Θ is parabolic derivable at $\lambda(X)$ for $\lambda'(X;H)$. Therefore, we have

$$\lambda''(X;H,Z) + p \in T^2_\Theta(\lambda(X), \lambda'(X;H)) \quad \text{with} \quad p := y - \lambda''(X;H,Z),$$

which can be equivalently expressed as $Z \in S_H(p)$ via the mapping S_H in (35). Appealing to Proposition 7 and the established outer Lipschitzian property in (37), we find a matrix $W \in S_H(0)$ such that $\|Z - W\| \le \|p\|$. This tells us that

$$\lambda''(X;H,W) \in T^2_\Theta(\lambda(X), \lambda'(X;H)).$$

Using the Chain Rule (42) leads us to $W \in T^2_C(X,H)$, and thus $T^2_C(X,H) \neq \emptyset$. This shows that C is parabolically derivable at X for H, and thus proves (a).

Turning into the proof of (b), observe first that Θ can be represented as the Constraint System (21). Adapting the estimate in (23) for the latter constraint system gives us the estimate

$$\text{dist}(x,\Theta) = \text{dist}(\text{Diag}\,(x), C) \quad \text{for all} \quad x \in \mathbf{R}^n.$$

This, together with twice differentiability of the mapping $x \longmapsto \text{Diag}\,(x)$ with $x \in \mathbf{R}^n$, allows us to conclude from (40) that for any $v \in T_\Theta(\lambda(X))$, we always have

$$w \in T^2_\Theta(\lambda(X), v) \Longleftrightarrow \text{Diag}\,(w) \in T^2_C(\text{Diag}\,(\lambda(X)),\ \text{Diag}\,(v)).$$

To justify (b), pick $v \in T_\Theta(\lambda(X))$. We are going to show that Θ is parabolically derivable at $\lambda(X)$ for v. According to Proposition 8, this will be ensured, provided that C is parabolically derivable at $\text{Diag}\,(\lambda(X))$ for $\text{Diag}\,(v)$. Because C is a spectral set, it is easy to see that

$$W \in T^2_C(\text{Diag}\,(\lambda(X)),\ \text{Diag}\,(v)) \Longleftrightarrow UWU^\top \in T^2_C(X, U\,\text{Diag}\,(v)U^\top), \tag{44}$$

where U is taken from (4). Moreover, it follows from $v \in T_\Theta(\lambda(X))$ and Proposition 5 that $\text{Diag}\,(v) \in T_C(\text{Diag}\,(\lambda(X)))$, which tells us that $U\,\text{Diag}\,(v)U^\top \in T_C(X)$. By assumption, we know that C is parabolically derivable at X for $U\,\text{Diag}\,(v)U^\top$. This, combined with (44), confirms that C is parabolically derivable at $\text{Diag}\,(\lambda(X))$ for $\text{Diag}\,(v)$. To justify this claim, take $W \in T^2_C(\text{Diag}\,(\lambda(X)),\ \text{Diag}\,(v))$. By (44), parabolic derivability of C at X for $UWU^\top$, and Proposition 3, we find $\varepsilon > 0$ such that for all $t \in [0,\varepsilon]$, the inclusion

$$X + tU\,\text{Diag}\,(v)U^\top + \frac{1}{2}t^2UWU^\top + o(t^2) \in C,$$

is satisfied. It follows from C being a spectral set and the latter inclusion that

$$\text{Diag}\,(\lambda(X)) + t\,\text{Diag}\,(v) + \frac{1}{2}t^2W + o(t^2) = U^\top XU + t\,\text{Diag}\,(v) + \frac{1}{2}t^2W + o(t^2) \in C \quad \text{for all} \ \ t \in [0,\varepsilon].$$

Because $W \in T^2_C(\text{Diag}\,(\lambda(X)),\ \text{Diag}\,(v))$ was taken arbitrarily, we conclude from Proposition 3 that C is parabolically derivable at $\text{Diag}\,(\lambda(X))$ for $\text{Diag}\,(v)$. Employing now Proposition 8 proves that Θ is parabolically derivable at $\lambda(X)$ for v and, hence, completes the proof. □

Example 3 (Second-Order Tangent Set to $\mathbf{S}^n_-$). In the framework of Example 1, we are going to calculate the second-order tangent set to $\mathbf{S}^n_-$ at $X \in \mathbf{S}^n_-$ for any $H \in T_{\mathbf{S}^n_-}(X)$. To this end, we deduce from Theorem 2 that

$$T^2_{\mathbf{S}^n_-}(X,H) = \{W \in \mathbf{S}^n \,|\, \lambda''(X;H,W) \in T^2_{\mathbf{R}^n_-}(\lambda(X), \lambda'(X;H))\}.$$

It follows from Rockafellar and Wets [25, proposition 13.12] that $\mathbf{R}^n_-$ is parabolically derivable at $\lambda(X)$ for $\lambda'(X;H)$, which together with Theorem 2 implies that $\mathbf{S}^n_-$ enjoys the same property at X for H. Moreover, we deduce from Rockafellar and Wets [25, proposition 13.12] that

$$T^2_{\mathbf{R}^n_-}(\lambda(X),\lambda'(X;H)) = T_{T_{\mathbf{R}^n_-}(\lambda(X))}(\lambda'(X;H)). \tag{45}$$

If $\mu_1 < 0$, we have $\lambda(X) \in \operatorname{int} \mathbf{R}^n_-$. This implies that $T_{\mathbf{R}^n_-}(\lambda(X)) = \mathbf{R}^n$, which together with (45) yields $T^2_{\mathbf{R}^n_-}(\lambda(X), \lambda'(X;H)) = \mathbf{R}^n$ and thus $T^2_{\mathbf{S}^n_-}(X,H) = \mathbf{S}^n$. Now, assume that $\mu_1 = 0$. According to Example 1, we have $T_{\mathbf{R}^n_-}(\lambda(X)) = \mathbf{R}^{|\alpha_1|}_- \times \mathbf{R}^{n-|\alpha_1|}$, where α_1 is defined by (7). To proceed, because $H \in T_{\mathbf{S}^n_-}(X)$, we need by (29) to consider two cases: (1) $\lambda_1(U^\top_{\alpha_1}HU_{\alpha_1}) < 0$; and (2) $\lambda_1(U^\top_{\alpha_1}HU_{\alpha_1}) = 0$. If the former holds, we obtain

$$T_{T_{\mathbf{R}^n_-}(\lambda(X))}(\lambda'(X;H)) = T_{\mathbf{R}^{|\alpha_1|}_-\times\mathbf{R}^{n-|\alpha_1|}}(\lambda'(X;H)) = \mathbf{R}^n,$$

which together with (45) brings us again to $T^2_{\mathbf{R}^n_-}(\lambda(X),\lambda'(X;H)) = \mathbf{R}^n$ and, thus, $T^2_{\mathbf{S}^n_-}(X,H) = \mathbf{S}^n$. If the latter holds, denote by $\eta^1_1 > \cdots > \eta^1_{\rho_1}$ the distinct eigenvalues of $U^\top_{\alpha_1}HU_{\alpha_1}$ and take the index set β^1_1 from (14). Recall that $|\beta^1_1| \le |\alpha_1|$. Using this, we obtain

$$T_{T_{\mathbf{R}^n_-}(\lambda(X))}(\lambda'(X;H)) = T_{\mathbf{R}^{|\alpha_1|}_-\times\mathbf{R}^{n-|\alpha_1|}}(\lambda'(X;H)) = \mathbf{R}^{|\beta^1_1|}_- \times \mathbf{R}^{n-|\beta^1_1|}.$$

This, combined with (16) and (45), leads us to

$$\begin{aligned} T^2_{\mathbf{S}^n_-}(X,H) &= \{W \in \mathbf{S}^n \,|\, \lambda''(X;H,W) \in \mathbf{R}^{|\beta^1_1|}_- \times \mathbf{R}^{n-|\beta^1_1|}\} \\ &= \{W \in \mathbf{S}^n \,|\, \lambda_1(R_{11}{}^\top(U^\top_{\alpha_1}(W - 2HX^\dagger H)U_{\alpha_1})R_{11}) \le 0\}, \end{aligned}$$

where $R_{11} = (Q_1)_{\beta^1_1}$ is taken from Proposition 2. We should point out that the second-order tangent set to $\mathbf{S}^n_-$ was calculated by finding the parabolic second-order directional derivative of the maximum eigenvalue function in Bonnans and Shapiro [3, p. 474]; see also Zhang et al. [28, p. 583] for a different derivation of this object.

In the next example, we obtain the second-order tangent set to the spectral set defined in (30). Note again that whereas obtaining such result by Bonnans and Shapiro [3, proposition 3.92] requires the Slater condition, our approach shows that no constraint qualification is needed for this purpose.

Example 4. Let C be the spectral set in (30) and $X \in C$. Given $H \in T_C(X)$, we aim to determine $T^2_C(X,H)$ using Theorem 2. We know from Example 2 that C has the spectral representation in (20) with the symmetric set Θ defined by (31). Moreover, we showed that Θ can be equivalently described as the constraint set in (32) with $\Phi(z) = (\sum_{i=1}^n z_i^k - 1, z)$ for all $z = (z_1,\ldots,z_n) \in \mathbf{R}^n$. We deduce from Rockafellar and Wets [25, proposition 13.13] that $w \in T^2_\Theta(\lambda(X),\lambda'(X;H))$ if and only if we have

$$\nabla\Phi(\lambda(X))w + \nabla^2\Phi(\lambda(X))(\lambda'(X;H),\lambda'(X;H)) \in T^2_{\{0\}\times\mathbf{R}^n_+}(\Phi(\lambda(X)),\nabla\Phi(\lambda(X))(\lambda'(X;H))). \tag{46}$$

Using the index set $I(\lambda(X))$ taken from Example 2, define the index set

$$I(\lambda(X),\lambda'(X;H)) := \{i \in I(\lambda(X)) \,|\, \lambda'_i(X;H) = 0\},$$

and conclude then from Rockafellar and Wets [25, proposition 13.12] that

$$\begin{aligned} T^2_{\{0\}\times\mathbf{R}^n_+}(\Phi(\lambda(X)),\nabla\Phi(\lambda(X))(\lambda'(X;H))) &= T_{T_{\{0\}\times\mathbf{R}^n_+}(\Phi(\lambda(X)))}(\nabla\Phi(\lambda(X))(\lambda'(X;H))) \\ &= \{(w_0,\ldots,w_n) \,|\, w_0 = 0,\, w_i \ge 0 \quad \text{for all} \;\; i \in I(\lambda(X),\lambda'(X;H))\}. \end{aligned}$$

This, coupled with (46), yields $(w_1,\ldots,w_n) \in T^2_\Theta(\lambda(X),\lambda'(X;H))$ if and only if $w_i \ge 0$ for all $i \in I(\lambda(X),\lambda'(X;H))$ and

$$\sum_{i=1}^n \lambda_i(X)^{k-1}w_i + (k-1)\sum_{i=1}^n \lambda_i(X)^{k-2}\lambda'_i(X;H)^2 = 0.$$

Appealing now to Theorem 2, we conclude that C is parabolically derivable at X for H and that $W \in T^2_C(X,H)$ if and only if $\lambda''_i(X;H,W) \ge 0$ for all $i \in I(\lambda(X),\lambda'(X;H))$ and

$$\sum_{i=1}^n \lambda_i(X)^{k-1}\lambda''_i(X;H,W) + (k-1)\sum_{i=1}^n \lambda_i(X)^{k-2}\lambda'_i(X;H)^2 = 0.$$

Note that when $k = 1$ for which C reduces to the spectahedron, the above equation simplifies as $\sum_{i=1}^n \lambda''_i(X;H,W) = 0$.

We proceed with characterizing parabolic epi-differentiability of spectral functions. We begin with recalling the concept of the parabolic subderivative, introduced by Ben-Tal and Zowe [2]. Let $f : \mathbf{X} \to \overline{\mathbf{R}}$ and let $\bar{x} \in \mathbf{X}$ with

$f(\bar{x})$ finite and $w \in \mathbf{X}$ 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

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

Recall from Rockafellar and Wets [25, definition 13.59] that f is called *parabolically epi-differentiable* at $\bar{x}$ for w if

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

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

$$\mathrm{d}^2 f(\bar{x})(w|z) = \lim_{k \to \infty} \frac{f\left(\bar{x} + t_k w + \frac{1}{2}t_k^2 z_k\right) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(w)}{\frac{1}{2}t_k^2}. \tag{47}$$

We say that f is parabolically epi-differentiable at $\bar{x}$ if it satisfies this condition at $\bar{x}$ for any $w \in \mathbf{X}$ where $\mathrm{d}f(\bar{x})(w)$ is finite. Note that the inclusion $\mathrm{dom}\ \mathrm{d}f(\bar{x}) \subset T_{\mathrm{dom} f}(\bar{x})$ always holds, and equality happens when, in addition, f is locally Lipschitz continuous around $\bar{x}$ relative to its domain; see Mohammadi and Sarabi [20, proposition 2.2]. A list of important functions, appearing in different classes of constrained and composite optimization problems, that are parabolically epi-differentiable at any points of their domains can be found in Mohammadi and Sarabi [20, example 4.7]. By definition, it is not hard to see that the inclusion

$$\mathrm{dom}\ \mathrm{d}^2 f(\bar{x})(w|\cdot) \subset T^2_{\mathrm{dom} f}(\bar{x}, w), \tag{48}$$

always holds for any $w \in T_{\mathrm{dom} f}(\bar{x})$. The following result, taken from Mohammadi and Sarabi [20, propositions 2.1 and 4.1], presents conditions under which we can ensure equality in the latter inclusion.

Proposition 9 (Properties of Parabolic Subderivatives). *Let $f : \mathbf{X} \to \overline{\mathbf{R}}$ be finite at $\bar{x}$, locally Lipschitz continuous around $\bar{x}$ relative to its domain, and parabolic epi-differentiable at $\bar{x}$ for $w \in T_{\mathrm{dom} f}(\bar{x})$. Then, the following properties hold.*

a. $\mathrm{dom}\ \mathrm{d}f(\bar{x}) = T_{\mathrm{dom} f}(\bar{x})$ *and* $\mathrm{dom}\ \mathrm{d}^2 f(\bar{x})(w|.) = T^2_{\mathrm{dom} f}(\bar{x}, w)$.

b. $\mathrm{dom}\, f$ *is parabolically derivable at* $\bar{x}$ *for* w.

The next result presents sufficient conditions under which spectral functions are parabolically epi-differentiable. Moreover, it achieves a useful formula for parabolic subderivatives of this class of functions.

Theorem 3 (Parabolic Subderivatives of Spectral Function). *Let $\theta : \mathbf{R}^n \to \overline{\mathbf{R}}$ be a symmetric function, which is locally Lipschitz continuous relative to its domain. Let $X \in \mathbf{S}^n$ with $(\theta \circ \lambda)(X)$ finite. Then, the following properties hold.*

a. *If $H \in T_{\mathrm{dom}\,(\theta\circ\lambda)}(X)$ and θ is parabolically epi-differentiable at $\lambda(X)$ for $\lambda'(X;H)$, then $\theta \circ \lambda$ is parabolically epi-differentiable at X for H, and its parabolic subderivative at X for H and its domain can be calculated, respectively, by*

$$\mathrm{d}^2(\theta \circ \lambda)(X)(H|W) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|\lambda''(X;H,W)), \tag{49}$$

and

$$\mathrm{dom}\ \mathrm{d}^2(\theta \circ \lambda)(X)(H|.) = T^2_{\mathrm{dom}\,(\theta\circ\lambda)}(X,H). \tag{50}$$

Moreover, if θ is lsc and convex, the parabolic subderivative $W \longmapsto \mathrm{d}^2(\theta \circ \lambda)(X)(H|W)$ is a convex function.

b. *If $\theta \circ \lambda$ is parabolically epi-differentiable at X, then θ is parabolically epi-differentiable at $\lambda(X)$.*

Proof. To justify (a), we proceed concurrently to show that $\theta \circ \lambda$ is parabolically epi-differentiable at X for H and that (49) and (47) hold for $\theta \circ \lambda$. To this end, set $g := \theta \circ \lambda$ and pick $W \in \mathbf{S}^n$ and proceed with considering two cases. Assume first that $W \notin T^2_{\mathrm{dom}\, g}(X,H)$. Employing the inclusion in (48) for g, we get $\mathrm{d}^2 g(X)(H|W) = \infty$. On the other hand, by (22) and Theorem 2, we obtain

$$T^2_{\mathrm{dom}\, g}(X,H) = \{W \in \mathbf{S}^n | \lambda''(X;H,W) \in T^2_{\mathrm{dom}\,\theta}(\lambda(X), \lambda'(X;H))\}. \tag{51}$$

This, combined with $W \notin T^2_{\mathrm{dom}\, g}(X,H)$, yields $\lambda''(X;H,W) \notin T^2_{\mathrm{dom}\,\theta}(\lambda(X), \lambda'(X;H))$. Observe from Corollary 2 and (22) that $H \in T_{\mathrm{dom}\, g}(X)$ amounts to $\lambda'(X;H) \in T_{\mathrm{dom}\,\theta}(\lambda(X))$. By Proposition 9(a), we arrive at

$$\mathrm{dom}\ \mathrm{d}^2\theta(\lambda(X))(\lambda(X;H)|\cdot) = T^2_{\mathrm{dom}\,\theta}(\lambda(X), \lambda'(X;H)). \tag{52}$$

Combining these tells us that $\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|\lambda''(X;H,W)) = \infty$, which, in turn, justifies (49) for every $W \notin T^2_{\mathrm{dom}\, g}(X,H)$. To verify (47) for g in this case, consider an arbitrary sequence $t_k \downarrow 0$, set $W_k := W$ for all $k \in \mathbf{N}$, and observe that

$$\infty = \mathrm{d}^2 g(X)(H|W) \le \liminf_{k\to\infty} \frac{g\big(X + t_k H + \frac{1}{2}t_k^2 W_k\big) - g(X) - t_k \mathrm{d}g(X)(W)}{\frac{1}{2}t_k^2}.$$

This clearly justifies (47) for all $W \notin T^2_{\mathrm{dom}\, g}(X,H)$.

Turning now to the case $W \in T^2_{\mathrm{dom}\, g}(X,H)$, we observe that because θ is parabolically epi-differentiable at $\lambda(X)$ for $\lambda'(X;H)$, Proposition 9(b) tells us that dom θ is parabolically derivable at $\lambda(X)$ for $\lambda'(X;H)$. We conclude from Theorem 2(a) that dom g is parabolically derivable at X for H. In particular, we have

$$T^2_{\mathrm{dom}\, g}(X,H) \ne \emptyset. \tag{53}$$

Pick now $W \in T^2_{\mathrm{dom}\, g}(X,H)$ and consider then an arbitrary sequence $t_k \downarrow 0$. Thus, by the definition of parabolic derivability, we find a sequence $W_k \to W$ as $k \to \infty$ such that

$$X_k := X + t_k H + \frac{1}{2}t_k^2 W_k = X + t_k H + \frac{1}{2}t_k^2 W + o(t_k^2) \in \mathrm{dom}\, g \quad \text{for all} \ \ k \in \mathbf{N}. \tag{54}$$

Moreover, because θ is parabolically epi-differentiable at $\lambda(X)$ for $\lambda'(X;H)$, we find a sequence $w_k \to w := \lambda''(X;H,W)$ such that

$$\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w) = \lim_{k\to\infty} \frac{\theta(\lambda(X) + t_k\lambda'(X;H) + \frac{1}{2}t_k^2 w_k) - \theta(\lambda(X)) - t_k \mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac{1}{2}t_k^2}.$$

It follows from (51) and $W \in T^2_{\mathrm{dom}\, g}(X,H)$ that $w \in T^2_{\mathrm{dom}\,\theta}(\lambda(X), \lambda'(X;H))$. Combining this with (52) tells us that $\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w) < \infty$. This implies that $y_k := \lambda(X) + t_k\lambda'(X;H) + \frac{1}{2}t_k^2 w_k \in \mathrm{dom}\,\theta$ for all k sufficiently large. Using this together with (26), (54), and (17), we obtain

$$\begin{aligned}
\mathrm{d}^2 g(X)(H|W) &\le \liminf_{k\to\infty} \frac{g\big(X + t_k H + \frac{1}{2}t_k^2 W_k\big) - g(X) - t_k \mathrm{d}g(X)(H)}{\frac{1}{2}t_k^2} \\
&\le \limsup_{k\to\infty} \frac{g\big(X + t_k H + \frac{1}{2}t_k^2 W_k\big) - g(X) - t_k \mathrm{d}g(X)(H)}{\frac{1}{2}t_k^2} \\
&= \limsup_{k\to\infty} \frac{\theta(\lambda(X_k)) - \theta(\lambda(X)) - t_k \mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac{1}{2}t_k^2} \\
&\le \limsup_{k\to\infty} \frac{\theta(y_k) - \theta(\lambda(X)) - t_k \mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac{1}{2}t_k^2} + \limsup_{k\to\infty} \frac{\theta(\lambda(X_k)) - \theta(y_k)}{\frac{1}{2}t_k^2} \\
&\le \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w) + \limsup_{k\to\infty}\ \ell \left\| \lambda''(X;H,W) - w_k + \frac{o(t_k^2)}{t_k^2} \right\| \\
&= \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w),
\end{aligned} \tag{55}$$

where $\ell \ge 0$ is a Lipschitz constant of θ around $\lambda(X)$ relative to its domain. On the other hand, for any sequence $t_k \downarrow 0$ and any sequence $W_k \to W$, we can always conclude from (17) and (26) that

$$\begin{aligned}
&\liminf_{k\to\infty} \frac{g\big(X + t_k H + \frac{1}{2}t_k^2 W_k\big) - g(X) - t_k \mathrm{d}g(X)(H)}{\frac{1}{2}t_k^2} \\
&= \liminf_{k\to\infty} \frac{\theta\big(\lambda(X) + t_k\lambda'(X;H) + \frac{1}{2}t_k^2\lambda''(X;H,W) + o(t_k^2)\big) - \theta(\lambda(X)) - t_k \mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac{1}{2}t_k^2} \\
&\ge \liminf_{\substack{t\downarrow 0 \\ w'\to w}} \frac{\theta\big(\lambda(X) + t\lambda'(X;H) + \frac{1}{2}t^2 w'\big) - \theta(\lambda(X)) - t\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac{1}{2}t^2} \\
&= \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w).
\end{aligned}$$

This clearly yields the inequality

$$\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|w) \le \mathrm{d}^2 g(X)(H|W) \quad \text{with} \ \ w = \lambda''(X;H,W).$$

Combining this and (55) implies that

$$d^2g(X)(H|W) = d^2\theta(\lambda(X))(\lambda'(X;H)|w), \tag{56}$$

and that

$$d^2g(X)(H|W) = \lim_{k\to\infty} \frac{g\big(X + t_kH + \frac{1}{2}t_k^2W_k\big) - g(X) - t_k dg(X)(H)}{\frac{1}{2}t_k^2},$$

which, in turn, prove both (49) and (47) for any $W \in T^2_{\mathrm{dom}\, g}(X,H)$. As argued above, we also have $d^2\theta(\lambda(X))(\lambda'(X;H)|w) < \infty$, which, together with (56), tells us that $d^2g(X)(H|W) < \infty$. This brings us to the inclusion

$$T^2_{\mathrm{dom}\, g}(X,H) \subset \mathrm{dom}\; d^2g(X)(H|\cdot).$$

Because the opposite inclusion always holds (see (48)), we arrive at (50). Combining this and (53) indicates that $\mathrm{dom}\; d^2g(X)(H|\cdot) \neq \emptyset$ and, hence, shows that g is parabolically epi-differentiable at X for H. Finally, assume that θ is lsc and convex. By Borwein and Lewis [4, corollary 5.2.3], the spectral function $g = \theta \circ \lambda$ is convex. According to Rockafellar and Wets [25, example 13.62], parabolic epi-differentiability of g at X for H amounts to parabolic derivability of epi g at $(X, g(X))$ for $(H, dg(X)(H))$ and

$$\mathrm{epi}\; d^2g(X)(H|.) = T^2_{\mathrm{epi}\, g}((X, g(X)), (H, dg(X)(H))).$$

Because g is convex, it follows from parabolic derivability of epi g at $(X, g(X))$ for $(H, dg(X)(H))$ that $T^2_{\mathrm{epi}\, g}((X, g(X)), (H, dg(X)(H)))$ is a convex set. The above equality then confirms that $W \longmapsto d^2(\theta \circ \lambda)(X)(H|W)$ is a convex function and, hence, completes the proof of (a).

Turning into the proof of (b), we conclude from (1) that the symmetric function θ satisfies (19), which means that θ can be represented as a composite function of g and the linear mapping $x \longmapsto \mathrm{Diag}\,(x)$ with $x \in \mathbf{R}^n$. We also deduce from the imposed assumption on θ and the inequality in (5) that g is locally Lipschitz continuous relative to its domain. Pick $v \in \mathrm{dom}\; d\theta(\lambda(X)) = T_{\mathrm{dom}\,\theta}(\lambda(X))$ and apply the chain rule in (25) to the representation (22) of dom θ to obtain $\mathrm{Diag}\,(v) \in T_{\mathrm{dom}\, g}(\mathrm{Diag}\,(\lambda(X)))$. To justify the parabolic epi-differentiability of θ at $\lambda(X)$ for v, we are going to use Mohammadi and Sarabi [20, theorem 4.4(iii)] by showing that g is parabolically epi-differentiable at $\mathrm{Diag}\,(\lambda(X))$ for $\mathrm{Diag}\,(v)$. To this end, it is not hard to see for any $U \in \mathbf{O}^n(X)$ that

$$dg(\mathrm{Diag}\,(\lambda(X)))(\mathrm{Diag}\,(v)) = dg(X)(U\,\mathrm{Diag}\,(v)U^\top). \tag{57}$$

Indeed, because g is orthogonally invariant, we get for any $U \in \mathbf{O}^n(X)$ that

$$\begin{aligned} dg(X)(U\,\mathrm{Diag}\,(v)U^\top) &= \liminf_{\substack{t\downarrow 0\\ W\to U\,\mathrm{Diag}\,(v)U^\top}} \frac{g(X+tW) - \theta(X))}{t} \\ &= \liminf_{\substack{t\downarrow 0\\ U^\top WU\to\ \mathrm{Diag}\,(v)}} \frac{g(\mathrm{Diag}\,(\lambda(X)) + tU^\top WU) - \theta(\mathrm{Diag}\,(\lambda(X)))}{t} \\ &\geq dg(\mathrm{Diag}\,(\lambda(X)))(\mathrm{Diag}\,(v)). \end{aligned}$$

A similar argument leads us to $dg(\mathrm{Diag}\,(\lambda(X)))(\mathrm{Diag}\,(v)) \leq dg(X)(U\,\mathrm{Diag}\,(v)U^\top)$ and, thus, proves (57). Similarly, we can show that

$$d^2g(\mathrm{Diag}\,(\lambda(X)))(\mathrm{Diag}\,(v)|W) = d^2g(X)(U\,\mathrm{Diag}\,(v)U^\top|UWU^\top), \quad W \in \mathbf{S}^n. \tag{58}$$

Because g is a spectral function, dom g is a spectral set. This and $\mathrm{Diag}\,(v) \in T_{\mathrm{dom}\, g}(\mathrm{Diag}\,(\lambda(X)))$ tell us that $U\,\mathrm{Diag}\,(v)U^\top \in T_{\mathrm{dom}\, g}(X)$. Because g is parabolically epi-differentiable at X for $U\,\mathrm{Diag}\,(v)U^\top$, the equality in (58) confirms that g enjoys the same property at $\mathrm{Diag}\,(\lambda(X))$ for $\mathrm{Diag}\,(v)$. Combining this, (23), and Mohammadi and Sarabi [20, theorem 4.4(iii)] shows that θ is parabolically epi-differentiable at $\lambda(X)$ for v and, hence, completes the proof. □

We close this section by revealing that the parabolic subderivative of spectral functions is symmetric with respect to a subset of $\mathbf{P}^n$, the set of all $n \times n$ permutation matrices. This plays a central role in the next section, when we are going to study parabolic regularity of spectral functions. To this end, recall from Remark 1 that if $\theta : \mathbf{R}^n \to \overline{\mathbf{R}}$ is a symmetric function and $X \in \mathbf{S}^n$ with $\theta(\lambda(X))$ finite, the subderivative function $d\theta(\lambda(X))$ is a symmetric function with respect to $\mathbf{P}^n_X$, which is a subset of $\mathbf{P}^n$ consisting of all $n \times n$ block diagonal matrices in the form $Q = \mathrm{Diag}\,(P_1, \ldots, P_r)$, where $P_m \in \mathbf{R}^{|\alpha_m|\times|\alpha_m|}$, $m = 1, \ldots, r$ is a permutation matrix with α_m taken from (7) and

r being the number of distinct eigenvalues of X. Consider now $H \in \mathbf{S}^n$ with $\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))$ finite. Take the orthogonal matrix U from (4) and $m \in \{1,\ldots,r\}$. Suppose that ρ_m is the number of distinct eigenvalues of $U_{\alpha_m}^\top H U_{\alpha_m}$ and pick then the index sets β_j^m for $j=1,\ldots,\rho_m$ from (14). Denote by $\mathbf{P}_{X,H}^n$ a subset of $\mathbf{P}_X^n$ consisting of all $n\times n$ matrices with representation Diag $(P_1,\ldots,P_r)$ such that for each $m=1,\ldots,r$, the $|\alpha_m|\times|\alpha_m|$ permutation matrix P_m has a block diagonal representation $P_m = \mathrm{Diag}\,(B_1^m,\ldots,B_{\rho_m}^m)$, where $B_j^m \in \mathbf{R}^{|\beta_j^m|\times|\beta_j^m|}$ is a permutation matrix for any $j=1,\ldots,\rho_m$. It is not hard to see that

$$Q\lambda(X) = \lambda(X) \quad \text{and} \quad Q\lambda'(X;H) = \lambda'(X;H) \quad \text{for any } Q \in \mathbf{P}_{X,H}^n. \tag{59}$$

Proposition 10. *Assume that $\theta:\mathbf{R}^n \to \overline{\mathbf{R}}$ is a symmetric function and $X \in \mathbf{S}^n$ with $\theta(\lambda(X))$ finite and that $H \in \mathbf{S}^n$ with $\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))$ finite. Then, for any $w \in \mathbf{R}^n$ and any permutation matrix $Q \in \mathbf{P}_{X,H}^n$, we have*

$$\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,Qw) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,w),$$

which means that the parabolic subderivative $w \mapsto \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,w)$ is symmetric with respect to $\mathbf{P}_{X,H}^n$.

Proof. Pick $w \in \mathbf{R}^n$ and $Q \in \mathbf{P}_{X,H}^n$. Because θ is symmetric, it follows from (59) that

$$\begin{aligned}
&\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,w)\\
&= \liminf_{\substack{t\downarrow 0\\ w'\to w}} \frac{\theta\big(\lambda(X)+t\lambda'(X;H)+\frac12 t^2 w'\big)-\theta(\lambda(X))-t\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac12 t^2}\\
&= \liminf_{\substack{t\downarrow 0\\ w'\to w}} \frac{\theta\big(\lambda(X)+t\lambda'(X;H)+\frac12 t^2 Qw'\big)-\theta(\lambda(X))-t\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac12 t^2}\\
&\ge \liminf_{\substack{t\downarrow 0\\ v\to Qw}} \frac{\theta\big(\lambda(X)+t\lambda'(X;H)+\frac12 t^2 v\big)-\theta(\lambda(X))-t\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))}{\frac12 t^2}\\
&= \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,Qw).
\end{aligned}$$

Similarly, using (59) for the matrix $Q^{-1} = \mathrm{Diag}\,(P_1^{-1},\ldots,P_r^{-1}) \in \mathbf{P}_{X,H}^n$, one can conclude that $\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,w) \le \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)\,|\,Qw)$, which justifies the claimed equality and, hence, ends the proof. □

5. Parabolic Regularity of Spectral Functions

This section is devoted to the study of parabolic regularity of spectral functions, whose central role in second-order variational analysis was revealed recently in Mohammadi et al. [21]. As demonstrated in Mohammadi et al. [21], parabolic regularity can be viewed as an important *second-order regularity* with remarkable consequences, among which we should highlight twice epi-differentiability of extended-real-valued functions. We begin with recalling the concepts of the second subderivative and parabolic regularity for functions, respectively. Given a function $f:\mathbf{X}\to\overline{\mathbf{R}}$ and $\bar{x}\in\mathbf{X}$ with $f(\bar{x})$ finite, define the parametric family of second-order difference quotients for f at $\bar{x}$ for $\bar{v}\in\mathbf{X}$ by

$$\Delta_t^2 f(\bar{x},\bar{v})(w) = \frac{f(\bar{x}+tw)-f(\bar{x})-t\langle \bar{v}, w\rangle}{\frac12 t^2} \qquad \text{with } w\in\mathbf{X},\ \ t>0.$$

The second subderivative of f at $\bar{x}$ for $\bar{v}$ is defined by

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

The importance of the second subderivative resides in the fact that it can characterize the quadratic growth condition for optimization problems; see Rockafellar and Wets [25, theorem 13.24]. So, it is crucial for many applications to calculate it in terms of the initial data of an optimization problem. This task was carried out for major classes of functions, including the convex piecewise linear-quadratic functions in the sense of Rockafellar and Wets [25, definition 10.20] in Rockafellar and Wets [25, proposition 13.9], the second-order/ice-cream cone in Mohammadi et al. [21, example 5.8], the cone of positive semidefinite symmetric matrices in Mohammadi and Sarabi [20, example 3.7], and the augmented Lagrangian of constrained optimization problems in Mohammadi et al. [21, theorem 8.3]. We are going to calculate it for the spectral function g in (1) when the symmetric function θ therein is convex.

Definition 1 (Parabolic Regularity). A convex function $f : \mathbf{X} \to \overline{\mathbf{R}}$ is parabolically regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if for any w such that $d^2 f(\bar{x},\bar{v})(w) < \infty$, there exist, among the sequences $t_k \downarrow 0$ and $w_k \to w$ with $\Delta^2_{t_k} f(\bar{x},\bar{v})(w_k) \to d^2 f(\bar{x},\bar{v})(w)$, those with the additional property that $\limsup_{k\to\infty} \|w_k - w\|/t_k < \infty$. We say that f is parabolically regular at $\bar{x}$ if it is parabolically regular at $\bar{x}$ for every $\bar{v} \in \partial f(\bar{x})$. A nonempty convex set $C \subset \mathbf{X}$ is said to be parabolically regular at $\bar{x}$ if the indicator function δ_C is parabolically regular at $\bar{x}$.

Parabolic regularity was introduced in Rockafellar and Wets [25, definition 13.65] for extended-real-valued functions, but was not scrutinized therein. It was shown in Mohammadi et al. [21] and Mohammadi and Sarabi [20] that polyhedral convex sets, the second-order/ice-cream cone, the cone of positive semidefinite symmetric matrices are parabolically regular. One can also find in Rockafellar and Wets [25, corollary 13.68] that convex piecewise linear-quadratic functions are parabolically regular. Recall that the critical cone of a convex function $f : \mathbf{X} \to \overline{\mathbf{R}}$ at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ is defined by

$$K_f(\bar{x},\bar{v}) = \{w \in \mathbf{X} \,|\, df(\bar{x})(w) = \langle \bar{v}, w\rangle\}.$$

When $f = \delta_C$, where C is a nonempty convex subset of $\mathbf{X}$, the critical cone of δ_C at $\bar{x}$ for $\bar{v}$ is denoted by $K_C(\bar{x},\bar{v})$. In this case, the above definition of the critical cone of a function boils down to the well-known concept of the critical cone of a set (see Dontchev and Rockafellar [12, p. 109]), namely, $K_C(\bar{x},\bar{v}) = T_C(\bar{x}) \cap [\bar{v}]^\perp$ because $d\delta_C(\bar{x}) = \delta_{T_C(\bar{x})}$.

The following result is a special case of a more general characterization of parabolic regularity from Mohammadi and Sarabi [20, proposition 3.6] and will be utilized in our approach in this section.

Proposition 11 (Characterization of Parabolic Regularity). *Assume that $f : \mathbf{X} \to \overline{\mathbf{R}}$ is convex, finite at $\bar{x} \in \mathbf{X}$, and $\bar{v} \in \partial f(\bar{x})$. Then, the following properties are equivalent.*

a. *f is parabolically regular at $\bar{x}$* for $\bar{v}$.

b. *For any $w \in K_f(\bar{x},\bar{v})$.*

$$d^2 f(\bar{x},\bar{v})(w) = \inf_{z\in\mathbf{X}}\{d^2 f(\bar{x})(w|z) - \langle z,\bar{v}\rangle\}. \tag{60}$$

c. *For any $w \in \mathrm{dom}\, d^2 f(\bar{x},\bar{v})$, there exists a $\bar{z} \in \mathrm{dom}\, d^2 f(\bar{x})(w|\cdot)$ such that*

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

Proof. The equivalence of (a) and (b) and the implication (a)⇒(c) were taken from Mohammadi and Sarabi [20, proposition 3.6]. To prove (c)⇒(b), take $w \in K_f(\bar{x},\bar{v})$. It follows from Rockafellar and Wets [25, proposition 13.64] that the inequality "≤" in (60) is always satisfied. To prove the opposite inequality, deduce first from the convexity of f that $d^2 f(\bar{x},\bar{v})$ is proper due to $d^2 f(\bar{x},\bar{v})(0) = 0$. By Rockafellar and Wets [25, proposition 13.5], the inclusion $\mathrm{dom}\, d^2 f(\bar{x},\bar{v}) \subset K_f(\bar{x},\bar{v})$ is satisfied. If $d^2 f(\bar{x},\bar{v})(w) = \infty$, the inequality "≥" in (60) trivially holds. Otherwise, $w \in \mathrm{dom}\, d^2 f(\bar{x},\bar{v})$, which, together with (c), proves the inequality "≥" in (60) and, hence, completes the proof. □

Our results so far required that the symmetric function θ in (1) be locally Lipschitz continuous with respect to its domain. In what follows, we need to assume further that θ is an lsc convex function. This assumption allows us to use the characterization of the subgradients of the spectral function g in (1), recorded in Proposition 6. We begin our analysis of the second subderivative of spectral functions by finding a lower estimate for it.

Proposition 12 (Lower Estimate for Second Subderivatives). *Assume that $g : \mathbf{S}^n \to \overline{\mathbf{R}}$ has the spectral representation in (1) and is lsc and convex and that $Y \in \partial g(X)$. Let $\mu_1 > \cdots > \mu_r$ be the distinct eigenvalues of X and $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$. Then, for any $H \in \mathbf{S}^n$, we have*

$$d^2 g(X,Y)(H) \ge d^2\theta(\lambda(X),\lambda(Y))(\lambda'(X;H)) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger H U_{\alpha_m}\rangle, \tag{61}$$

where α_m, $m = 1,\ldots,r$, are defined in (7).

Proof. Let $H\in\mathbf{S}^n$ and pick sequences $H_k\to H$ and $t_k\downarrow 0$. Setting $\Delta_{t_k}\lambda(X)(H_k):=(\lambda(X+t_kH_k)-\lambda(X))/t_k$, we get

$$\begin{aligned}\Delta^2_{t_k}g(X,Y)(H_k)&=\frac{\theta(\lambda(X+t_kH_k))-\theta(\lambda(X))-t_k\langle Y,H_k\rangle}{\frac{1}{2}t_k^2}\\&=\frac{\theta(\lambda(X)+t_k\Delta_{t_k}\lambda(X)(H_k))-\theta(\lambda(X))-t_k\langle\lambda(Y),\Delta_{t_k}\lambda(X)(H_k)\rangle}{\frac{1}{2}t_k^2}\\&\quad+\frac{\langle\lambda(Y),\Delta_{t_k}\lambda(X)(H_k)\rangle-\langle Y,H_k\rangle}{\frac{1}{2}t_k}\\&=\Delta^2_{t_k}\theta(\lambda(X),\lambda(Y))(\Delta_{t_k}\lambda(X)(H_k))+\frac{\langle\lambda(Y),\Delta_{t_k}\lambda(X)(H_k)\rangle-\langle Y,H_k\rangle}{\frac{1}{2}t_k}.\end{aligned}$$

It follows from $Y=U\Lambda(Y)U^\top$ that

$$\langle Y,H_k\rangle=\langle U\Lambda(Y)U^\top,H_k\rangle=\langle\Lambda(Y),U^\top H_kU\rangle=\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U^\top_{\alpha_m}H_kU_{\alpha_m}\rangle. \tag{62}$$

On the other hand, it results from (10) and Fan's inequality that

$$\begin{aligned}\langle\lambda(Y),\Delta_{t_k}\lambda(X)(H_k)\rangle&=\sum_{m=1}^r\sum_{j\in\alpha_m}\frac{\lambda_j(Y)(\lambda_j(X+t_kH_k)-\lambda_j(X))}{t_k}\\&=\sum_{m=1}^r\sum_{j\in\alpha_m}\lambda_j(Y)\lambda_{\ell_j}(U^\top_{\alpha_m}H_kU_{\alpha_m}+t_kU^\top_{\alpha_m}H_k(\mu_mI-X)^\dagger H_kU_{\alpha_m})+O(t_k^2)\\&\geq\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U^\top_{\alpha_m}H_kU_{\alpha_m}+t_kU^\top_{\alpha_m}H_k(\mu_mI-X)^\dagger H_kU_{\alpha_m}\rangle+O(t_k^2).\end{aligned}$$

Combining this with (62) brings us to

$$\frac{\langle\lambda(Y),\Delta_{t_k}\lambda(X)(H_k)\rangle-\langle Y,H_k\rangle}{\frac{1}{2}t_k}\geq 2\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U^\top_{\alpha_m}H_k(\mu_mI-X)^\dagger H_kU_{\alpha_m}\rangle+O(t_k).$$

This leads us to the estimate

$$\begin{aligned}\Delta^2_{t_k}g(X,Y)(H_k)&\geq\Delta^2_{t_k}\theta(\lambda(X),\lambda(Y))(\Delta_{t_k}\lambda(X)(H_k))\\&\quad+2\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U^\top_{\alpha_m}H_k(\mu_mI-X)^\dagger H_kU_{\alpha_m}\rangle+O(t_k),\end{aligned}$$

which, in turn, clearly justifies the lower estimate in (61) for the second subderivative of g at X for Y because $\Delta_{t_k}\lambda(X)(H_k)\to\lambda'(X;H)$ as $k\to\infty$. □

We proceed with a result about the critical cone of spectral functions.

Proposition 13 (Critical Cone of Spectral Functions). *Assume that $g:\mathbf{S}^n\to\overline{\mathbf{R}}$ has the spectral representation in (1) and is lsc and convex and that $Y\in\partial g(X)$. Let $\mu_1>\cdots>\mu_r$ be the distinct eigenvalues of X. Then, we have $H\in K_g(X,Y)$ if and only if $\lambda'(X;H)\in K_\theta(\lambda(X),\lambda(Y))$ and the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U^\top_{\alpha_m}HU_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m=1,\dots,r$ with α_m taken from (7) and $U\in\mathbf{O}^n(X)\cap\mathbf{O}^n(Y)$.*

Proof. By Borwein and Lewis [4, corollary 5.2.3], the symmetric function θ in (1) is lsc and convex. Thus, we find $U\in\mathbf{O}^n(X)\cap\mathbf{O}^n(Y)$ and get $\lambda(Y)\in\partial\theta(\lambda(X))$ due to Proposition 6. Pick $H\in K_g(X,Y)$ and deduce from the definition of the critical cone that $\mathrm{d}g(X)(H)=\langle Y,H\rangle$. We then conclude from $Y=U\Lambda(Y)U^\top$ with $\Lambda(Y)=\mathrm{Diag}\,(\lambda(Y))$ and Fan's inequality that

$$\begin{aligned}\langle Y,H\rangle=\langle\Lambda(Y),U^\top HU\rangle&=\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U^\top_{\alpha_m}HU_{\alpha_m}\rangle\\&\leq\sum_{m=1}^r\langle\lambda(Y)_{\alpha_m},\lambda(U^\top_{\alpha_m}HU_{\alpha_m})\rangle=\langle\lambda(Y),\lambda'(X;H)\rangle\\&\leq\mathrm{d}\theta(\lambda(X))(\lambda'(X;H))=\mathrm{d}g(X)(H),\end{aligned} \tag{63}$$

where the last inequality results from $\lambda(Y) \in \partial\theta(\lambda(X))$, Rockafellar and Wets [25, exercise 8.4], and the convexity of θ and where the last equality comes from (26). These relationships clearly imply that $\mathrm{d}\theta(\lambda(X))(\lambda'(X;H)) = \langle \lambda(Y), \lambda'(X;H)\rangle$, meaning that $\lambda'(X;H) \in K_\theta(\lambda(X), \lambda(Y))$, and that $\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H U_{\alpha_m}\rangle = \langle \lambda(Y)_{\alpha_m}, \lambda(U_{\alpha_m}^\top H U_{\alpha_m})\rangle$ for any $m = 1, \ldots, r$. By Fan's inequality, these equalities are equivalent to saying that the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m = 1, \ldots, r$.

To prove the opposite claim, assume $\lambda'(X;H) \in K_\theta(\lambda(X), \lambda(Y))$ and the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m = 1, \ldots, r$. The latter tells us via Fan's inequality that $\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H U_{\alpha_m}\rangle = \langle \lambda(Y)_{\alpha_m}, \lambda(U_{\alpha_m}^\top H U_{\alpha_m})\rangle$ for any $m = 1, \ldots, r$. Moreover, the former yields $\mathrm{d}\theta(\lambda(X))(\lambda'(X;H)) = \langle \lambda(Y), \lambda'(X;H)\rangle$. Taking these into account demonstrates that both inequalities in (63) are indeed equalities. This leads us to $\mathrm{d}g(X)(H) = \langle Y, H\rangle$, which implies that $H \in K_g(X, Y)$. □

The characterization of the critical cone of spectral functions, obtained above, is a generalization of a similar result, established recently in Cui and Ding [6, proposition 4] for the spectral function g from (1) when the symmetric function θ therein is a polyhedral function, meaning a function that its epigraph is a polyhedral convex set. To obtain a full characterization of the critical cone of spectral functions, we should know when the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ in Proposition 13 have a simultaneous ordered spectral decomposition because the critical cone $K_\theta(\lambda(X), \lambda(Y))$ can often be calculated rather easily. Although this remains an open question for now and will be a subject of our future research, we show by an example below the possible role that the condition on $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ is playing in the calculation of the critical cone of spectral functions. Indeed, if the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition, it is possible to show that the matrix $U_{\alpha_m}^\top H U_{\alpha_m}$ has a block diagonal structure; see (68) below and the discussion afterward to see why this can happen in the case $g = \delta_{\mathbf{S}_+^n}$.

Example 5. Set $g = \delta_{\mathbf{S}_+^n}$. Clearly, g is a spectral function satisfying (1) with $\theta = \delta_{\mathbf{R}_+^n}$. Take $Y \in N_{\mathbf{S}_+^n}(X)$ and observe from Proposition 6 that $Y = U\,\mathrm{Diag}\,(\lambda(Y))U^\top$, where $\lambda(Y) \in N_{\mathbf{R}_+^n}(\lambda(X))$ and $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$. Our goal is to calculate $K_{\mathbf{S}_+^n}(X, Y)$ using the characterization of this cone from Proposition 13. Take $H \in K_{\mathbf{S}_+^n}(X, Y)$, assume that $\mu_1 > \cdots > \mu_r$ are the distinct eigenvalues of X, and pick the constants α_m for any $m = 1, \ldots, r$ from (7). By Proposition 13, we conclude that $\lambda'(X;H) \in K_{\mathbf{R}_+^n}(\lambda(X), \lambda(Y))$ and that the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m = 1, \ldots, r$. The former is equivalent to the conditions

$$w \in T_{\mathbf{R}_+^n}(\lambda(X)) \quad \text{and} \quad \langle \lambda(Y), w\rangle = 0 \quad \text{with} \quad w = (w_1, \ldots, w_n) := \lambda'(X;H). \tag{64}$$

Moreover, the inclusion $\lambda(Y) \in N_{\mathbf{R}_+^n}(\lambda(X))$ amounts to

$$\sum_{i=1}^{n} \lambda_i(X)\lambda_i(Y) = 0, \quad \lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X)) \in \mathbf{R}_+^n, \quad \lambda(Y) = (\lambda_1(Y), \ldots, \lambda_n(Y)) \in \mathbf{R}_-^n. \tag{65}$$

Define the index sets

$$\kappa_X := \{i \in \{1, \ldots, n\} \,|\, \lambda_i(X) > 0\} \quad \text{and} \quad \tau_X := \{i \in \{1, \ldots, n\} \,|\, \lambda_i(X) = 0\},$$

and

$$\kappa_Y := \{i \in \{1, \ldots, n\} \,|\, \lambda_i(Y) = 0\} \quad \text{and} \quad \tau_Y := \{i \in \{1, \ldots, n\} \,|\, \lambda_i(Y) < 0\}.$$

It is easy to see from the definition of the index set α_r in (7) that $\tau_X = \alpha_r$ and from (65) that

$$\kappa_X \subset \kappa_Y \quad \text{and} \quad \tau_Y \subset \tau_X, \tag{66}$$

and to conclude from (64) and (65) that

$$w_i \in \begin{cases} \mathbf{R} & \text{if } i \in \kappa_X, \\ \mathbf{R}_+ & \text{if } i \in \tau_X \setminus \tau_Y, \\ \{0\} & \text{if } i \in \tau_Y. \end{cases} \tag{67}$$

Take $m \in \{1, \ldots, r-1\}$ and observe from the first inclusion in (66) that $\Lambda(Y)_{\alpha_m\alpha_m} = 0$. In this case, we will not benefit further from the fact that the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition. It remains to take a closer look into the case $m = r$. Because $\Lambda(Y)_{\alpha_r\alpha_r}$ and $U_{\alpha_r}^\top H U_{\alpha_r}$ have a simultaneous ordered spectral decomposition, we find an orthogonal matrix $Q_r \in \mathbf{O}^{|\alpha_r|}$ such that

$$\Lambda(Y)_{\alpha_r\alpha_r} = Q_r \Lambda(Y)_{\alpha_r\alpha_r} Q_r^\top \quad \text{and} \quad U_{\alpha_r}^\top H U_{\alpha_r} = Q_r \Lambda(U_{\alpha_r}^\top H U_{\alpha_r}) Q_r^\top. \tag{68}$$

Similar to (7), define the index sets $\{\rho_r^\nu\}_{\nu=1}^{\ell}$, where ℓ in the number of distinct eigenvalues of $\Lambda(Y)_{\alpha_r\alpha_r}$, by

$$\begin{cases} \lambda_i(Y) = \lambda_j(Y) & \text{if } \ i,j \in \rho_r^\nu \\ \lambda_i(Y) > \lambda_j(Y) & \text{if } \ i \in \rho_r^\nu,\ j \in \rho_r^k \ \text{ with } \ \nu < k. \end{cases}$$

Thus, because $\rho_r^j \subset \alpha_r = \tau_X$ for any $j \in \{1, \ldots, \ell\}$, we deduce from the third condition in (65) that

$$\rho_r^1 = \tau_X \cap \kappa_Y \quad \text{and} \quad \bigcup_{j=2}^{\ell} \rho_r^j = \tau_X \cap \tau_Y = \tau_Y. \tag{69}$$

It is not hard to see from the first equality in (68) (see Ding [10, proposition 2.4] for more detail) that Q_r has a block diagonal representation as

$$Q_r = \text{Diag}\,(Q_r^1, \ldots, Q_r^\ell) = \begin{pmatrix} Q_r^1 & 0 & \cdots & 0 \\ 0 & Q_r^2 & \ddots & \vdots \\ \vdots & \ddots & Q_r^{\ell-1} & 0 \\ 0 & \cdots & 0 & Q_r^\ell \end{pmatrix} \quad \text{with } \ Q_r^j \in \mathbf{O}^{|\rho_r^j|},\ \ j = 1, \ldots, \ell.$$

This, coupled with the second equality in (68), implies that $U_{\alpha_r}^\top H U_{\alpha_r}$ has a similar block diagonal representation as

$$U_{\alpha_r}^\top H U_{\alpha_r} = \text{Diag}\,(H_r^1, \ldots, H_r^\ell) = \begin{pmatrix} H_r^1 & 0 & \cdots & 0 \\ 0 & H_r^2 & \ddots & \vdots \\ \vdots & \ddots & H_r^{\ell-1} & 0 \\ 0 & \cdots & 0 & H_r^\ell \end{pmatrix} \quad \text{with } \ H_r^j \in \mathbf{S}^{|\rho_r^j|},\ \ j = 1, \ldots, \ell.$$

A direct calculation shows that $H_r^j = U_{\rho_r^j}^\top H U_{\rho_r^j}$ for any $j = 1, \ldots, \ell$. Moreover, it follows from (67), (69), and the definition of w_i from (64) that

$$U_{\rho_r^1}^\top H U_{\rho_r^1} \in \mathbf{S}_+^{|\rho_r^1|} \quad \text{and} \quad U_{\rho_r^j}^\top H U_{\rho_r^j} = 0 \quad \text{for any } \ j = 2, \ldots, \ell,$$

which, in turn, leads us to the representation

$$U_{\alpha_r}^\top H U_{\alpha_r} = \left(\begin{array}{c|c} U_{\rho_r^1}^\top H U_{\rho_r^1} & 0 \\ \hline 0 & 0 \end{array}\right). \tag{70}$$

This gives us the inclusion

$$K_{\mathbf{S}_+^n}(X, Y) \subset \{H \in \mathbf{S}^n \,|\, U_{\rho_r^1}^\top H U_{\rho_r^1} \in \mathbf{S}_+^{|\rho_r^1|},\ U_{\tau_Y}^\top H U_{\tau_Y} = 0,\ U_{\rho_r^1}^\top H U_{\tau_Y} = 0\}. \tag{71}$$

We claim now that the inclusion above becomes equality. To prove it, take a matrix H from the right-hand side of the above inclusion. To justify $H \in K_{\mathbf{S}_+^n}(X, Y)$, we first show that $\lambda'(X; H) \in K_{\mathbf{R}_+^n}(\lambda(X), \lambda(Y))$, which is equivalent to proving (64). By the selection of H, the components of the vector $w = \lambda'(X; H)$ enjoy the properties in (67), because $\alpha_r = \tau_X$ and $\rho_r^1 = \tau_X \cap \kappa_Y = \tau_X \setminus \tau_Y$ due to (66). Now, it is not hard to see that w satisfies all the conditions in (64), confirming the inclusion $\lambda'(X; H) \in K_{\mathbf{R}_+^n}(\lambda(X), \lambda(Y))$. To finish the proof, it suffices, according to Proposition 13, to demonstrate that the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m = 1, \ldots, r$. Take $m \in \{1, \ldots, r-1\}$ and observe that if $i \in \alpha_m \subset \kappa_X$, it follows from the first inclusion in (66) that $\lambda_i(Y) = 0$. This implies that $\Lambda(Y)_{\alpha_m\alpha_m} = 0$ for any such an m, which, in turn, tells us that $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top H U_{\alpha_m}$ have a simultaneous ordered spectral decomposition. It remains to consider the case $m = r$. We know from the selection of H that $U_{\alpha_r}^\top H U_{\alpha_r}$ has the representation in (70). According to (69), the diagonal matrix $\Lambda(Y)_{\alpha_r\alpha_r}$ has a similar block structure as (70), given by

$$\Lambda(Y)_{\alpha_r\alpha_r} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \Lambda(Y)_{\gamma\gamma} \end{array}\right) \quad \text{with } \ \gamma := \alpha_r \setminus \rho_r^1.$$

Taking into account this representation and (70) tells us that $\Lambda(Y)_{\alpha_r\alpha_r}$ and $U_{\alpha_r}^\top H U_{\alpha_r}$ have a simultaneous ordered spectral decomposition and, hence, finishes the proof of the opposite inclusion in (71). We should add here that the same description as (71) for $K_{\mathbf{S}_+^n}(X,Y)$ was obtained in Chan and Sun [5] using a different approach and without appealing to a characterization of the latter cone obtained in Proposition 13.

Note that the analysis above for the case of $\delta_{\mathbf{S}_+^n}$ clearly illustrates the essential role that the simultaneous ordered spectral decompositions of $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^\top HU_{\alpha_m}$ play in finding the critical cone of spectral functions. To shed more light into the role of the later condition, consider the case $n=3$, $X=\text{Diag}\,(1,0,0)$ and $Y=\text{Diag}\,(0,0,-1)$. In this case, we get $r=2$, $\alpha_1=\{1\}$, and $\alpha_2=\{2,3\}$. Moreover, we have $\kappa_X=\{1\}$, $\tau_X=\{2,3\}$, $\kappa_Y=\{1,2\}$, $\tau_Y=\{3\}$, and $\rho_r^1=\{2\}$. According to (71), the symmetric matrix H belongs to $K_{\mathbf{S}_+^n}(X,Y)$ if and only if it has a representation of the form

$$\begin{pmatrix} \star & \star & \star \\ \star & a & 0 \\ \star & 0 & 0 \end{pmatrix},$$

where $a\in\mathbf{R}_+$ and the $\star$ positions can be filled with any real number. This shows the matrix

$$H=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & -1/2 & 0 \end{pmatrix},$$

doesn't belong to $K_{\mathbf{S}_+^3}(X,Y)$ because H_{23} and H_{32} are not zero. To elaborate more on why this happens, observe first that $U:=I\in\mathbf{O}^3(X)\cap\mathbf{O}^3(Y)$. Thus, we have

$$\lambda'(X;H)=(\lambda(U_{\alpha_1}^\top HU_{\alpha_1}),\lambda(U_{\alpha_2}^\top HU_{\alpha_2}))=(\lambda(H_{\alpha_1\alpha_1}),\lambda(H_{\alpha_2\alpha_2}))=\left(1,\frac{1+\sqrt{5}}{4},\frac{1-\sqrt{5}}{4}\right),$$

which clearly belongs to $K_{\mathbf{R}_+^3}(\lambda(X),\lambda(Y))$. This can be justified via the equivalent description of $K_{\mathbf{R}_+^3}(\lambda(X),\lambda(Y))$ in (64). However, it is possible to demonstrate that $\Lambda(Y)_{\alpha_2\alpha_2}$ and $U_{\alpha_2}^\top HU_{\alpha_2}$ don't have a simultaneous ordered spectral decomposition by showing that

$$\langle\Lambda(Y)_{\alpha_2\alpha_2},U_{\alpha_2}^\top HU_{\alpha_2}\rangle\neq\langle\lambda(\Lambda(Y)_{\alpha_2\alpha_2}),\lambda(U_{\alpha_2}^\top HU_{\alpha_2})\rangle. \tag{72}$$

This, indeed, results from the fact that Fan's inequality in (6) becomes equality if and only if the matrices therein have a simultaneous ordered spectral decomposition. To prove (72), we deduce from $\Lambda(Y)_{\alpha_2\alpha_2}=\text{Diag}\,(0,-1)$ and $U_{\alpha_2}^\top HU_{\alpha_2}=H_{\alpha_2\alpha_2}$ that

$$\langle\Lambda(Y)_{\alpha_2\alpha_2},U_{\alpha_2}^\top HU_{\alpha_2}\rangle=\text{tr}\left(\begin{pmatrix} 0 & 0 \\ 0 & -1\end{pmatrix}\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 0\end{pmatrix}\right)=0.$$

On the other hand, we have

$$\langle\lambda(\Lambda(Y)_{\alpha_2\alpha_2}),\lambda(U_{\alpha_2}^\top HU_{\alpha_2})\rangle=\frac{\sqrt{5}-1}{4},$$

which confirms (72) and, hence, tells us that the main reason for $H\notin K_{\mathbf{S}_+^3}(X,Y)$ is the failure of ensuring a simultaneous ordered spectral decomposition for $\Lambda(Y)_{\alpha_2\alpha_2}$ and $U_{\alpha_2}^\top HU_{\alpha_2}$.

As mentioned before, we can partition any vector $p\in\mathbf{R}^n$ into $(p_{\alpha_1},\ldots,p_{\alpha_r})$ with α_m, $m=1,\ldots,r$, taken from (7). Pick $m\in\{1,\ldots,r\}$ and recall from (14) that the index set $\alpha_m=\cup_{i=1}^{\rho_m}\beta_i^m$. This allows us to partition further p_{α_m} into $(y_{\beta_1^m},\ldots,y_{\beta_{\rho_m}^m})$, where $y_{\beta_i^m}\in\mathbf{R}^{|\beta_i^m|}$ for any $i=1,\ldots,\rho_m$. In summary, we can equivalently write p as

$$(y_{\beta_1^1},\ldots,y_{\beta_{\rho_1}^1},\ldots,y_{\beta_1^r},\ldots,y_{\beta_{\rho_r}^r}), \tag{73}$$

where r, taken from (7), and ρ_m, taken from (14), stand for the number of distinct eigenvalues of X and $U_{\alpha_m}^\top HU_{\alpha_m}$, respectively. Thus, the representation of p in (73) is associated with the permutation matrices in $\mathbf{P}_{X,H}^n$, defined prior to Proposition 10, as a subset of $\mathbf{P}^n$. In fact, any permutation matrix $Q\in\mathbf{P}_{X,H}^n$ has a representation of the form $\text{Diag}\,(B_1^1,\ldots,B_{\rho_1}^1,\ldots,B_1^r,\ldots,B_{\rho_r}^r)$, where $B_j^m\in\mathbf{R}^{|\beta_j^m|\times|\beta_j^m|}$ is a permutation matrix for any $j\in\{1,\ldots,\rho_m\}$ and $m\in\{1,\ldots,r\}$. Denote by $\mathbf{R}_\downarrow^n$ the set of all vectors $(x_1,\ldots,x_n)$ such that $x_1\geq\cdots\geq x_n$.

Proposition 14. *Assume that the spectral function $g=\theta\circ\lambda$ in (1) is lsc and convex and that $Y\in\partial g(X)$ and $H\in K_g(X,Y)$. If θ is parabolically regular at $\lambda(X)$ for $\lambda(Y)$, then the following properties hold.*

a. *There exists* $z \in \mathbf{R}^n$, *which has a representation in the form in* (73) *with* $y_{\beta_i^m} \in \mathbf{R}_{\downarrow}^{|\beta_i^m|}$ *for any* $i \in \{1, \ldots, \rho_m\}$ *and* $m \in \{1, \ldots, r\}$, *satisfying*

$$\mathrm{d}^2\theta(\lambda(X), \lambda(Y))(\lambda'(X;H)) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|z) - \langle \lambda(Y), z\rangle. \tag{74}$$

b. *There exists a matrix* $\widehat{W} \in \mathbf{S}^n$ *such that* $\lambda''(X;H, \widehat{W}) = z$, *where z comes from* (*a*).

Proof. We deduce from $Y \in \partial g(X)$ and Proposition 6 that $\lambda(Y) \in \partial\theta(\lambda(X))$. Also, it follows from $H \in K_g(X,Y)$ and Proposition 13 that $\lambda'(X;H) \in K_\theta(\lambda(X), \lambda(Y))$. Employing now Proposition 11 ensures the existence of $p \in \mathbf{R}^n$ satisfying (74). As explained above, any such a vector p has a representation in the form of (73) with $y_{\beta_i^m} \in \mathbf{R}^{|\beta_i^m|}$ for any $i \in \{1, \ldots, \rho_m\}$ and $m \in \{1, \ldots, r\}$. We are going to show that we can find a vector p with the representation in (73) such that $y_{\beta_i^m} \in \mathbf{R}_{\downarrow}^{|\beta_i^m|}$ for any $i \in \{1, \ldots, \rho_m\}$ and $m \in \{1, \ldots, r\}$, meaning the components of each $y_{\beta_i^m}$ have nonincreasing order. To this end, pick $m \in \{1, \ldots, r\}$ and $i \in \{1, \ldots, \rho_m\}$ and choose then a $|\beta_i^m| \times |\beta_i^m|$ permutation matrix B_i^m such that $q_{\beta_i^m} := B_i^m y_{\beta_i^m} \in \mathbf{R}_{\downarrow}^{|\beta_i^m|}$. Set $Q := \mathrm{Diag}\,(B_1^1, \ldots, B_{\rho_1}^1, \ldots, B_1^r, \ldots, B_{\rho_r}^r)$ and observe that $Q \in \mathbf{P}_{X,H}^n$. Moreover, let

$$z := (q_{\beta_1^1}, \ldots, q_{\beta_{\rho_1}^1}, \ldots, q_{\beta_1^r}, \ldots, q_{\beta_{\rho_r}^r}). \tag{75}$$

Clearly, we have $z = Qp$. It follows from Proposition 10 that

$$\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|z) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|p).$$

We claim now that $\langle \lambda(Y), p\rangle \le \langle \lambda(Y), z\rangle$. To justify it, suppose that

$$(\lambda(Y)_{\beta_1^1}, \ldots, \lambda(Y)_{\beta_{\rho_1}^1}, \ldots, \lambda(Y)_{\beta_1^r}, \ldots, \lambda(Y)_{\beta_{\rho_r}^r}),$$

is a partition of the vector $\lambda(Y)$ corresponding to (73). Note that $\lambda(Y)_{\beta_i^m} \in \mathbf{R}_{\downarrow}^{|\beta_i^m|}$ for any $i \in \{1, \ldots, \rho_m\}$ and $m \in \{1, \ldots, r\}$. Thus, we get

$$\langle \lambda(Y), p\rangle = \sum_{m=1}^{r}\sum_{i=1}^{\rho_m} \langle \lambda(Y)_{\beta_i^m}, y_{\beta_i^m}\rangle \le \sum_{m=1}^{r}\sum_{i=1}^{\rho_m} \langle \lambda(Y)_{\beta_i^m}, q_{\beta_i^m}\rangle = \langle \lambda(Y), z\rangle,$$

where the inequality is a consequence of the Hardy-Littlewood-Pólya inequality (cf. Borwein and Lewis [4, proposition 1.2.4]). Set $\varphi(x) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|x) - \langle \lambda(Y), x\rangle$ for any $x \in \mathbf{R}^n$ and observe from Proposition 11 that p is a minimizer of φ. But we showed above that $\varphi(z) \le \varphi(p)$, which tells us that z is also a minimizer of φ. Thus, we arrive at $\varphi(z) = \varphi(p)$, which implies that (74) holds for z. This proves (a).

Turning now to the proof of (b), pick the vector z from (75). We can equivalently write via the index sets α_m, $m = 1, \ldots, r$, from (7) that

$$z = (z_{\alpha_1}, \ldots, z_{\alpha_r}) \quad \text{with} \quad z_{\alpha_m} = (q_{\beta_1^m}, \ldots, q_{\beta_{\rho_m}^m}) \in \mathbf{R}^{|\alpha_m|} \quad \text{for all} \;\; m \in \{1, \ldots, r\}. \tag{76}$$

Take the $|\alpha_m| \times |\alpha_m|$ matrix Q_m, $m = 1, \ldots, r$, from (13) and consider the $n \times n$ block diagonal matrix

$$A = \mathrm{Diag}\,(Q_1\, \mathrm{Diag}\,(z_{\alpha_1})Q_1^\top, \ldots, Q_r\, \mathrm{Diag}\,(z_{\alpha_r})Q_r^\top). \tag{77}$$

We claim that there exists a matrix $\widehat{W} \in \mathbf{S}^n$ such that for any $m = 1, \ldots, r$ the relationship

$$U_{\alpha_m}^\top\, \widehat{W} U_{\alpha_m} = U_{\alpha_m}^\top (2H(X - \mu_m I)^\dagger H + UAU^\top)U_{\alpha_m}, \tag{78}$$

holds, where $\mu_1 > \cdots > \mu_r$ are the distinct eigenvalues of X and $U \in \mathbf{O}^n(X)$. Indeed, to find such a matrix $\widehat{W}$, let $W \in \mathbf{S}^n$ and set $\widehat{W} = UWU^\top$ in the above equality. This, coupled with (9), leads us to

$$W_{\alpha_m\alpha_m} = U_{\alpha_m}^\top UWU^\top U_{\alpha_m} = U_{\alpha_m}^\top\, \widehat{W} U_{\alpha_m} = U_{\alpha_m}^\top (2H(X - \mu_m I)^\dagger H + UAU^\top)U_{\alpha_m},$$

for all $m = 1, \ldots, r$. Define the matrix W as the block diagonal matrix $\mathrm{Diag}\,(W_{\alpha_1\alpha_1}, \ldots, W_{\alpha_r\alpha_r})$, from which we can obtain the claimed matrix $\widehat{W}$. Suppose now that $i \in \{1, \ldots, n\}$. By (15), there are $m \in \{1, \ldots, r\}$ and $j \in \{1, \ldots, \rho_m\}$ such that $i \in \alpha_m$ and $\ell_i \in \beta_j^m$. According to (16), we have

$$\begin{aligned}
\lambda_i''(X;H, \widehat{W}) &= \lambda_{\ell_i'}((Q_m)_{\beta_j^m}^\top (U_{\alpha_m}^\top(\widehat{W} + 2H(\mu_m I - X)^\dagger H)U_{\alpha_m})(Q_m)_{\beta_j^m}) \\
&= \lambda_{\ell_i'}((Q_m)_{\beta_j^m}^\top (U_{\alpha_m}^\top UAU^\top U_{\alpha_m})(Q_m)_{\beta_j^m}) \\
&= \lambda_{\ell_i'}((Q_m)_{\beta_j^m}^\top Q_m\, \mathrm{Diag}\,(z_{\alpha_m})Q_m^\top (Q_m)_{\beta_j^m}) \\
&= \lambda_{\ell_i'}(\mathrm{Diag}\,(q_{\beta_j^m})),
\end{aligned} \tag{79}$$

where the last two equalities result from (9). Consider now a partition of $\lambda''(X;H,\widehat{W})$ corresponding to (75) as

$$(\eta_{\beta_1^1},\ldots,\eta_{\beta_{\rho_1}^1},\ldots,\eta_{\beta_1^r},\ldots,\eta_{\beta_{\rho_r}^r}).$$

Thus, it follows from (79) and the definition ℓ_i' that

$$\eta_{\beta_j^m}=\lambda(\mathrm{Diag}\,(q_{\beta_j^m}))=q_{\beta_j^m},$$

for any $j\in\{1,\ldots,\rho_m\}$ and $m\in\{1,\ldots,r\}$ because $q_{\beta_j^m}\in\mathbf{R}_{\downarrow}^{|\beta_j^m|}$. This confirms that $\lambda''(X;H,\widehat{W})=z$ and, hence, completes the proof of (b). □

We are now ready to characterize parabolic regularity of spectral functions. We begin with the following result, in which we provide a sufficient condition to calculate the domain of the second subderivative of spectral functions.

Proposition 15. *Assume that the spectral function $g=\theta\circ\lambda$ in (1) is convex and that $Y\in\partial g(X)$. Then, we have* $\mathrm{dom}\,\mathrm{d}^2g(X,Y)\subset K_g(X,Y)$. *Equality holds if, in addition, g is parabolically epi-differentiable at X for any $H\in K_g(X,Y)$.*

Proof. The claimed inclusion results from Mohammadi and Sarabi [20, proposition 2.1(ii)–(iii)]. To establish the second claim, it follows from parabolic epi-differentiability of g at X for any $H\in K_g(X,Y)$ that $\mathrm{dom}\,\mathrm{d}^2g(X)(H|.)\neq\emptyset$. This, coupled with Mohammadi and Sarabi [20, proposition 3.4], confirms that $\mathrm{dom}\,\mathrm{d}^2g(X,Y)=K_g(X,Y)$ and hence completes the proof. □

Note that the assumption of parabolic epi-differentiability of g in the result above can be ensured via Theorem 3(a) by parabolic epi-differentiability of θ. An important class of functions for which this assumption automatically is satisfied is polyhedral functions; see Rockafellar and Wets [25, exercise 13.61]. This class of functions allows us to cover many examples of spectral functions, which are important for applications. We should add that polyhedral functions that are symmetric were characterized in Cui and Ding [6, proposition 1].

Our next result presents a characterization of parabolic regularity of spectral functions.

Theorem 4 (Parabolic Regularity of Spectral Functions). *Assume that the spectral function $g=\theta\circ\lambda$ in (1) is locally Lipschitz continuous with respect to its domain, lsc, and convex. Let $\mu_1>\cdots>\mu_r$ be the distinct eigenvalues of X. Then, the following properties hold.*

a. *If $Y\in\partial g(X)$ and θ is parabolically regular at $\lambda(X)$ for $\lambda(Y)$ and parabolically epi-differentiable at $\lambda(X)$, then g is parabolically regular at X for Y, and for any $H\in K_g(X,Y)$, we have*

$$\mathrm{d}^2g(X,Y)(H)=\mathrm{d}^2\theta(\lambda(X),\lambda(Y))(\lambda'(X;H))+2\sum_{m=1}^{r}\langle\Lambda(Y)_{\alpha_m\alpha_m},U_{\alpha_m}^{\top}H(\mu_mI-X)^{\dagger}HU_{\alpha_m}\rangle,$$

where α_m, $m=1,\ldots,r$, come from (7), $U\in\mathbf{O}^n(X)\cap\mathbf{O}^n(Y)$, and $\Lambda(Y)=\mathrm{Diag}\,(\lambda(Y))$.

b. *If g is parabolically epi-differentiable and parabolically regular at X, then θ is parabolically regular at $\lambda(X)$.*

Proof. We begin with the proof of (a). To justify (a), it suffices by Proposition 11 to show that for any $H\in K_g(X,Y)$, we have

$$\mathrm{d}^2g(X,Y)(H)=\inf_{W\in\mathbf{S}^n}\{\mathrm{d}^2g(X)(H|W)-\langle Y,W\rangle\}. \tag{80}$$

To this end, pick $H\in K_g(X,Y)$ and deduce from Rockafellar and Wets [25, proposition 13.64] and (49), respectively, that

$$\begin{aligned}\mathrm{d}^2g(X,Y)(H)&\le\inf_{W\in\mathbf{S}^n}\{\mathrm{d}^2g(X)(H|W)-\langle Y,W\rangle\}\\&=\inf_{W\in\mathbf{S}^n}\{\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|\lambda''(X;H,W))-\langle Y,W\rangle\}.\end{aligned} \tag{81}$$

Because $H\in K_g(X,Y)$, it results from Proposition 13 that the matrices $\Lambda(Y)_{\alpha_m\alpha_m}$ and $U_{\alpha_m}^{\top}HU_{\alpha_m}$ have a simultaneous ordered spectral decomposition for any $m=1,\ldots,r$. This means that there are matrices $\widehat{Q}_m\in\mathbf{O}^{|\alpha_m|}(\Lambda(Y)_{\alpha_m\alpha_m})\cap\mathbf{O}^{|\alpha_m|}(U_{\alpha_m}^{\top}HU_{\alpha_m})$, $m=1,\ldots,r$, such that

$$\Lambda(Y)_{\alpha_m\alpha_m}=\widehat{Q}_m\Lambda(Y)_{\alpha_m\alpha_m}\widehat{Q}_m^{\top}\quad\text{and}\quad U_{\alpha_m}^{\top}HU_{\alpha_m}=\widehat{Q}_m\Lambda(U_{\alpha_m}^{\top}HU_{\alpha_m})\widehat{Q}_m^{\top}. \tag{82}$$

Replace the matrices Q_m in the definition of the matrix A in (77) with $\widehat{Q}_m$, $m=1,\ldots,r$, and observe that the same conclusion can be achieved as the one in Proposition 14(b) for the updated matrix A. In fact, the matrices $\widehat{Q}_m$

enjoy all the properties of Q_m together with the relationships in (82), which are important for our argument below. Because θ is parabolically regular at $\lambda(X)$ for $\lambda(Y)$, we conclude from Proposition 14(a) that there is $z \in \mathbf{R}^n$ with a representation in the form in (75) with $q_{\beta_i^m} \in \mathbf{R}_{\downarrow}^{|\beta_i^m|}$ for any $i \in \{1, \ldots, \rho_m\}$ and $m \in \{1, \ldots, r\}$, satisfying (74). According to Proposition 14(b), there exists a matrix $\widehat{W} \in \mathbf{S}^n$ such that $\lambda''(X; H, \widehat{W}) = z$. Employing now (74) and (78), and $Y = U\Lambda(Y)U^\top$ and using a similar argument as (62), we arrive at

$$
\begin{aligned}
&\mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|\lambda''(X;H,\widehat{W})) - \langle Y, \widehat{W}\rangle \\
&= \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|z) - \sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top \widehat{W} U_{\alpha_m}\rangle \\
&= \mathrm{d}^2\theta(\lambda(X), \lambda(Y))(\lambda'(X;H)) + \langle \lambda(Y), z\rangle \\
&\quad - \sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top (2H(X - \mu_m I)^\dagger H + UAU^\top)U_{\alpha_m}\rangle \\
&= \mathrm{d}^2\theta(\lambda(X), \lambda(Y))(\lambda'(X;H)) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top (H(\mu_m I - X)^\dagger H)U_{\alpha_m}\rangle \\
&\quad + \langle \lambda(Y), z\rangle - \sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top (UAU^\top)U_{\alpha_m}\rangle. \qquad (83)
\end{aligned}
$$

By the definition of A in (77), the equality in (9), and the representation of the vector z in (76), we obtain

$$
\begin{aligned}
\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top (UAU^\top)U_{\alpha_m}\rangle &= \sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, \widehat{Q}_m \,\mathrm{Diag}\,(z_{\alpha_m})\widehat{Q}_m^\top\rangle \\
&= \sum_{m=1}^{r}\langle \widehat{Q}_m^\top \Lambda(Y)_{\alpha_m\alpha_m}\widehat{Q}_m, \,\mathrm{Diag}\,(z_{\alpha_m})\rangle \\
&= \sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, \,\mathrm{Diag}\,(z_{\alpha_m})\rangle = \langle \lambda(Y), z\rangle,
\end{aligned}
$$

where the penultimate equality results from the first relationship in (82) and the last one is a consequence of $\Lambda(Y) = \mathrm{Diag}\,(\lambda(Y))$. This, coupled with (61), (81), and (83), brings us to

$$
\begin{aligned}
&\mathrm{d}^2\theta(\lambda(X), \lambda(Y))(\lambda'(X;H)) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger H U_{\alpha_m}\rangle \\
&\le \mathrm{d}^2 g(X, Y)(H) \\
&\le \inf_{W\in\mathbf{S}^n}\{\mathrm{d}^2 g(X)(H|W) - \langle Y, W\rangle\} \\
&\le \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)|\lambda''(X;H,\widehat{W})) - \langle Y, \widehat{W}\rangle \\
&= \mathrm{d}^2\theta(\lambda(X), \lambda(Y))(\lambda'(X;H)) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger H U_{\alpha_m}\rangle.
\end{aligned}
$$

These relationships clearly justify (80) and, hence, imply that g is parabolically regular at X for Y. Moreover, they confirm the claimed formula for the second subderivative of g at X for Y for any $H \in K_g(X, Y)$ and, hence, complete the proof of (a).

Turning into the proof of (b), we need to show that θ is parabolically regular at $\lambda(X)$ for any $v \in \partial\theta(\lambda(X))$. To justify it, pick $v \in \partial\theta(\lambda(X))$ and $U \in \mathbf{O}^n(X)$ and deduce from Lewis [18, theorem 6] that $U\,\mathrm{Diag}\,(v)U^\top \in \partial g(X)$. Because g is convex and orthogonally invariant, the latter yields $\mathrm{Diag}\,(v) \in \partial g(\Lambda(X))$, where $\Lambda(X) = \mathrm{Diag}\,(\lambda(X))$. We claim that g is parabolically regular at $\Lambda(X)$ for $\mathrm{Diag}\,(v)$. To this end, set $Z := U\,\mathrm{Diag}\,(v)U^\top$ and take $H \in K_g(\Lambda(X), \mathrm{Diag}\,(v))$. We conclude from (57) and the definition of the critical cone that $UHU^\top \in K_g(X, Z)$. Thus, it also follows from parabolic regularity of g at X for Z and Proposition 11 that there is $W \in \mathbf{S}^n$ such that

$$
\mathrm{d}^2 g(X, Z)(UHU^\top) = \mathrm{d}^2 g(X)(UHU^\top|W) - \langle Z, W\rangle. \qquad (84)
$$

Because g is orthogonally invariant, we get

$$\begin{aligned}
d^2g(X,Z)(UHU^\top) &= \liminf_{\substack{t\downarrow 0\\ H'\to UHU^\top}} \frac{g(X+tH')-g(X)-\langle Z,H'\rangle}{\frac{1}{2}t^2}\\
&= \liminf_{\substack{t\downarrow 0\\ H'\to UHU^\top}} \frac{g(\Lambda(X)+tU^\top H'U)-g(\Lambda(X))-\langle U^\top ZU,U^\top H'U\rangle}{\frac{1}{2}t^2}\\
&\ge d^2g(\Lambda(X),\ \text{Diag}\,(v))(H).
\end{aligned}$$

Similarly, we can show that $d^2g(X,Z)(UHU^\top)\le d^2g(\Lambda(X),\ \text{Diag}\,(v))(H)$, which leads us to $d^2g(X,Z)(UHU^\top)=d^2g(\Lambda(X),\ \text{Diag}\,(v))(H)$. It follows from (58) that

$$d^2g(X)(UHU^\top\,|\,W)=d^2g(\Lambda(X))(H\,|\,U^\top WU).$$

We also have $\langle Z,W\rangle=\langle\text{Diag}\,(v),U^\top WU\rangle$. Combining these with (84) brings us to

$$d^2g(\Lambda(X),\ \text{Diag}\,(v))(H)=d^2g(\Lambda(X))(H\,|\,U^\top WU)-\langle\text{Diag}\,(v),U^\top WU\rangle.$$

Because $H\in K_g(\Lambda(X),\ \text{Diag}\,(v))$ was taken arbitrarily and because $\text{dom}\, d^2g(\Lambda(X),\ \text{Diag}\,(v))\subset K_g(\Lambda(X),\ \text{Diag}\,(v))$ is satisfied because of Proposition 15, we conclude from Proposition 11 that g is parabolically regular at $\Lambda(X)$ for $\text{Diag}\,(v)$. Recall that the symmetric function θ satisfies (19), which means that θ can be represented as $\theta=g\circ F$ with $F(x):=\text{Diag}\,(x)$ for all $x\in\mathbf{R}^n$. We also deduce from the imposed assumption on θ and (5) that g is locally Lipschitz continuous relative to its domain. According to Mohammadi et al. [22, theorem 3.6], we get $\partial\theta(\lambda(X))=\nabla F(\lambda(X))^*\partial g(\Lambda(X))$. It is not hard to see that for any $x\in\mathbf{R}^n$, $\nabla F(x)$ is a linear operator from $\mathbf{R}^n$ into $\mathbf{S}^n$, defined by $\nabla F(x)(y)=\sum_{i=1}^n y_iE_{ii}$ for any $y=(y_1,\dots,y_n)\in\mathbf{R}^n$, where E_{ii}, $i=1,\dots,n$, are the $n\times n$ matrix with $(i,\,i)$ entry equal to one and elsewhere equal to zero. This tells us that the adjoint operator $\nabla F(x)^*:\mathbf{S}^n\to\mathbf{R}^n$ has a representation in the form $\nabla F(x)^*B=(\text{tr}\,(E_{11}B),\dots,\ \text{tr}\,(E_{nn}B))$ for any $B\in\mathbf{S}^n$; see Beck [1, example 1.8] for more details. Remember that $v\in\partial\theta(\lambda(X))$ and $\text{Diag}\,(v)\in\partial g(\Lambda(X))$. Thus, we have $\nabla F(\lambda(X))^*\ \text{Diag}\,(v)=v$. On the other hand, it follows from the proof of Theorem 3(b) that parabolic epi-differentiability of g at X for $UHU^\top\in\text{dom}\, dg(X)$ implies that of g at $\Lambda(X)$ for H. Combining these and Mohammadi and Sarabi [20, theorem 5.4] tells us that θ is parabolically regular at $\lambda(X)$ for v and, hence, completes the proof of (b). □

Given the spectral function g in (1), it was shown in Cui et al. [7, proposition 10] that if θ is $\mathcal{C}^2$-cone reducible in the sense of Cui et al. [7, definition 6], then g enjoys the same property. Note that $\mathcal{C}^2$-cone reducibility of θ is strictly stronger assumption than parabolic regularity of this function, utilized in Theorem 4, as shown in Mohammadi et al. [21, theorem 6.2 and example 6.4]. Note also we showed in Theorem 4(b) that parabolic regularity of g yields that of θ; such a result was not achieved for $\mathcal{C}^2$-cone reducibility in Cui et al. [7].

In many important applications of the spectral function g in (1), the symmetric function θ is a polyhedral function. In this case, all the assumptions in Theorem 4 are satisfied automatically. Furthermore, the second subderivative of g has a simple representation as demonstrated below.

Corollary 3. *Assume that $g:\mathbf{S}^n\to\overline{\mathbf{R}}$ has the spectral representation in (1) with the symmetric function θ being polyhedral. If $\mu_1>\cdots>\mu_r$ are the distinct eigenvalues of X and $Y\in\partial g(X)$, then g is parabolically regular at X for Y, and for any $H\in\mathbf{S}^n$, we have*

$$d^2g(X,Y)(H)=\delta_{K_g(X,Y)}(H)+2\sum_{m=1}^r\langle\Lambda(Y)_{\alpha_m\alpha_m},U_{\alpha_m}^\top H(\mu_mI-X)^\dagger HU_{\alpha_m}\rangle,$$

where α_m, $m=1,\dots,r$, come from (7), $U\in\mathbf{O}^n(X)\cap\mathbf{O}^n(Y)$, and $\Lambda(Y)=\text{Diag}\,(\lambda(Y))$.

Proof. It follows from Rockafellar and Wets [25, exercise 13.61] and Mohammadi and Sarabi [20, example 3.2], respectively, that θ is parabolically epi-differentiable and parabolically regular at $\lambda(X)$. By Theorem 4(a), the spectral function g is parabolically regular at X for Y. Moreover, we know from Proposition 15 that $\text{dom}\, d^2g(X,Y)=K_g(X,Y)$. Take $H\in K_g(X,Y)$ and observe from Proposition 13 that $\lambda'(X;H)\in K_\theta(\lambda(X),\lambda(Y))$. Thus, we obtain from Rockafellar and Wets [25, proposition 13.9] that

$$d^2\theta(\lambda(X),\lambda(Y))(\lambda'(X;H))=\delta_{K_\theta(\lambda(X),\lambda(Y))}(\lambda'(X;H))=0.$$

Employing now Theorem 4(a) tells us that

$$d^2 g(X,Y)(H) = 2\sum_{m=1}^{r} \langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger H U_{\alpha_m}\rangle \quad \text{for all} \quad H \in K_g(X,Y).$$

On the other hand, if $H \notin K_g(X,Y)$, we have $d^2 g(X,Y)(H) = \infty$ because $\operatorname{dom} d^2 g(X,Y) = K_g(X,Y)$. This proves the claimed formula for the second subderivative of g at X for Y. □

Note that the conjugate function of the parabolic subderivative of the spectral function g in (1) with θ therein being polyhedral was recently calculated in Cui and Ding [6, propositions 6 and 10] by dividing a polyhedral function into two parts. In general, such a result gives us an upper bound for the second subderivative (cf. Rockafellar and Wets [25, proposition 13.64]). According to Proposition 11, parabolic regularity is, indeed, equivalent to saying that the latter conjugate function coincides with the second subderivative. We should add that parabolic regularity of g was not discussed in Cui and Ding [6], and so Corollary 4 can't be derived from the aforementioned results in Cui and Ding [6].

We continue to apply the formula of the second subderivative, obtained in Corollary 3, for two important examples of spectral functions and show how one can simplify the established formula for the second subderivative in these cases.

Example 6.

a. Assume that $g : \mathbf{S}^n \to \mathbf{R}$ is defined by $g(X) = \lambda_{\max}(X)$, where $\lambda_{\max}$ stands for the maximum eigenvalue of X. g is a spectral function and satisfies the representation (1) and $\theta(x) = \max\{x_1, \ldots, x_n\}$ with $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. Take $Y \in \partial g(X)$ and observe from Proposition 6 that $Y = U \operatorname{Diag}(\lambda(Y))U^\top$, where $\lambda(Y) \in \partial\theta(\lambda(X))$ and $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$. Recall from (7) that $\alpha_1 = \{i \in \{1, \ldots, n\} \mid \lambda_i(X) = \lambda_{\max}(X)\}$. It follows from Rockafellar and Wets [25, exercise 8.31] that

$$\partial\theta(\lambda(X)) = \left\{ (t_1, \ldots, t_n) \,\middle|\, \sum_{i=1}^{n} t_i = 1,\ t_i \ge 0 \text{ for all } i \in \alpha_1,\ t_i = 0 \ \text{ otherwise} \right\}.$$

Taking into consideration the formula for the second subderivative from Corollary 3 and the notation therein and the description of $\partial g(\lambda(X))$, we conclude that $\Lambda(Y)_{\alpha_m\alpha_m} = 0$ for all $m \ge 2$. Moreover, we have

$$Y = U \operatorname{Diag}(\lambda(Y))U^\top = \sum_{i=1}^{n} \lambda_i(Y)U_iU_i^T = \sum_{i\in\alpha_1} \lambda_i(Y)U_iU_i^T = U_{\alpha_1}\Lambda(Y)_{\alpha_1\alpha_1}U_{\alpha_1}^\top.$$

Combining this and Corollary 3, we obtain for any $H \in \mathbf{S}^n$ that

$$\begin{aligned} d^2 g(X,Y)(H) &= \delta_{K_g(X,Y)}(H) + 2\langle \Lambda(Y)_{\alpha_1\alpha_1}, U_{\alpha_1}^\top H(\mu_1 I - X)^\dagger H U_{\alpha_1}\rangle \\ &= \delta_{K_g(X,Y)}(H) + 2\langle Y, H(\mu_1 I - X)^\dagger H\rangle. \end{aligned}$$

This is the same formula, obtained in Torki [26, theorem 2.1], for the second subderivative of the first leading eigenvalue of a symmetric matrix. Note that it was proven in Mohammadi and Sarabi [20, example 3.3] that all leading eigenvalues of a symmetric matrix is parabolically regular. Also, one can find their second subderivatives in Torki [26, theorem 2.1]. Because the leading eigenvalues, except the first one, which is the maximum eigenvalue, are not convex, Theorem 4 and Corollary 3 can't be utilized to cover them. That requires to extend the established theory in this section for subdifferentially regular functions in the sense of Rockafellar and Wets [25, definition 7.25], a task that we leave for our future research.

b. Suppose that $g = \delta_{\mathbf{S}^n_-}$. As shown in Example 1, g is a spectral function satisfying (1) with $\theta = \delta_{\mathbf{R}^n_-}$. Take $Y \in N_{\mathbf{S}^n_-}(X)$ and observe from Proposition 6 that $Y = U \operatorname{Diag}(\lambda(Y))U^\top$, where $\lambda(Y) \in N_{\mathbf{R}^n_-}(\lambda(X))$ and $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$. If $\mu_1 = \lambda_1(X) < 0$, it follows from $X \in \operatorname{int} \mathbf{S}^n_-$ that g is twice differentiable and $d^2 g(X,Y)(H) = 0$ for any $H \in \mathbf{S}^n$. Assume now that $\mu_1 = \lambda_1(X) = 0$. Recall from (7) that $\alpha_1 = \{i \in \{1, \ldots, n\} \mid \lambda_i(X) = \mu_1\}$. Thus, we obtain

$$N_{\mathbf{R}^n_-}(\lambda(X)) = \{(t_1, \ldots, t_n) \mid t_i \ge 0 \text{ for all } i \in \alpha_1,\ t_i = 0 \ \text{ otherwise}\}.$$

Arguing similarly to (a) leads us to

$$d^2 \delta_{\mathbf{S}^n_-}(X,Y)(H) = \delta_{K_{\mathbf{S}^n_-}(X,Y)}(H) - 2\langle Y, HX^\dagger H\rangle \quad \text{for all } H \in \mathbf{S}^n.$$

This formula was obtained previously in Mohammadi and Sarabi [20, example 3.7] via a different approach. Similarly, we can show that if $Y \in N_{\mathbf{S}^n_+}(X)$, the second subderivative of $\delta_{\mathbf{S}^n_+}$ at X for Y can be calculated by

$$\mathrm{d}^2\delta_{\mathbf{S}^n_+}(X,Y)(H) = \delta_{K_{\mathbf{S}^n_+}(X,Y)}(H) - 2\langle Y, HX^\dagger H\rangle \quad \text{for all } H \in \mathbf{S}^n.$$

The second subderivative can be utilized to establish second-order optimality conditions for different classes of optimization problems. Doing so often requires obtaining a chain rule for the second subderivative, a task carried out in Theorem 4 and Corollary 3. Given a twice differentiable function $\varphi : \mathbf{S}^n \to \mathbf{R}$ and a spectral function g, consider the optimization problem

$$\text{minimize } \varphi(X) + g(X) \quad \text{subject to} \quad X \in \mathbf{S}^n. \tag{85}$$

Below, we present a result in which second-order optimality conditions for this optimization problem are established. For simplicity, we are going to assume that g has the assumed representation in Corollary 3, but one can easily extend it for any g satisfying the assumptions in Theorem 4.

Theorem 5. *Assume that X is a feasible solution to (85), where the spectral function g has the representation (1) with θ therein being a polyhedral function. If $-\nabla\varphi(X) \in \partial g(X)$, then the following second-order optimality conditions for (85) are satisfied.*

a. *If X is a local minimizer of (85), then the second-order necessary condition*

$$\nabla^2\varphi(X)(H,H) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger HU_{\alpha_m}\rangle \geq 0,$$

holds for all $H \in K_g(X, -\nabla\varphi(X))$.

b. *The validity of the second-order condition*

$$\nabla^2\varphi(X)(H,H) + 2\sum_{m=1}^{r}\langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger HU_{\alpha_m}\rangle > 0,$$

for all $H \in K_g(X, -\nabla\varphi(X))$ amounts to the existence of the constants $\ell \geq 0$ and $\varepsilon > 0$ for which the quadratic growth condition

$$\varphi(X') + g(X') \geq \varphi(X) + g(X) + \frac{\ell}{2}\|X' - X\|^2 \quad \textit{for all} \quad X' \in \mathbf{B}_\varepsilon(X),$$

is satisfied.

Proof. It follows from Rockafellar and Wets [25, exercise 13.18] that

$$\mathrm{d}^2(\varphi + g)(X,0)(H) = \nabla^2\varphi(X)(H,H) + \mathrm{d}^2 g(X, -\nabla\varphi(X))(H),$$

for any $H \in \mathbf{S}^n$. Both claims in (a) and (b) then result immediately from Rockafellar and Wets [25, theorem 13.24] and Corollary 3. □

Our next application is to provide sufficient conditions for twice epi-differentiability of spectral functions, a concept with important consequences in second-order variational analysis and parametric optimization. Recall from Rockafellar and Wets [25, definition 13.6] that a function $f : \mathbf{X} \to \overline{\mathbf{R}}$ is said to be twice epi-differentiable at $\bar{x}$ for $\bar{v} \in \mathbf{X}$, with $f(\bar{x})$ finite, if the sets epi $\Delta_t^2 f(\bar{x},\bar{v})$ converge to epi $\mathrm{d}^2 f(\bar{x},\bar{v})$ as $t \downarrow 0$ in the sense of set convergence from Rockafellar and Wets [25, definition 4.1], where "epi " stands for the epigraph of a function. This can be equivalently described via Rockafellar and Wets [25, proposition 7.2] that for every sequence $t_k \downarrow 0$ and every $w \in \mathbf{X}$, there exists a sequence $w_k \to w$ such that

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

Twice epi-differentiability is a geometric notion of second-order approximation for extended-real-valued functions and was defined by Rockafellar [23]. Its central role in second-order variational analysis, parametric optimization, and numerical algorithms has been demonstrated in Rockafellar and Wets [25], Mohammadi et al. [21], Mohammadi and Sarabi [20], and Hang and Sarabi [16]. It was observed recently in Mohammadi and Sarabi [20, corollary 5.5] that parabolic regularity of certain composite functions yields their twice epi-differentiability. A similar conclusion can be drawn for spectral functions as demonstrated below.

Corollary 4 (Twice Epi-Differentiability of Spectral Functions). *Assume that the spectral function $g = \theta \circ \lambda$ in (1) is locally Lipschitz continuous with respect to its domain, lsc, and convex. If $Y \in \partial g(X)$ and θ is parabolically regular at $\lambda(X)$ for $\lambda(Y)$ and parabolically epi-differentiable at $\lambda(X)$, then g is twice epi-differentiable at X for Y.*

Proof. The claimed twice epi-differentiability of g at X for Y results directly from Mohammadi and Sarabi [20, theorem 3.8] and Theorem 4. □

Twice epi-differentiability of leading eigenvalues and the sum of largest eigenvalues of a symmetric matrix was established in Torki [26, theorem 2.1] using a different approach. Corollary 4 goes far beyond the framework in Torki [26] to achieve twice epi-differentiability of spectral functions. We, however, can't get this property for all leading eigenvalues, except the first one, from Corollary 4 because these spectral functions are not convex. As explained in Example 6(a), this can be accomplished if the established theory in this section is generalized for subdifferentially regular functions. Note also that a characterization of directionally differentiability of the proximal mapping of spectral functions can be found in Ding et al. [11, theorem 3]. Recall from Beck [1, theorem 7.18] that if the spectral function g in (1) is lsc and convex, its proximal mapping can be calculated by

$$\operatorname{prox}_g(X) := \operatorname*{arg\,min}_{W \in \mathbf{S}^n} \left\{ g(W) + \frac{1}{2}\|W - X\|^2 \right\} = U \operatorname{Diag}\,(\operatorname{prox}_\theta(\lambda(X)))U^\top,$$

where $U \in \mathbf{O}^n(X)$. It follows from Ding et al. [11, theorem 3] that prox_g is directionally differentiable at X if and only if $\operatorname{prox}_\theta$ is directionally differentiable at $\lambda(X)$. It also follows from Rockafellar and Wets [25, exercise 13.45] that twice epi-differentiability of g at X for Y amounts to directional differentiability of prox_g at $X + Y$. Combining these, we can conclude that g is twice epi-differentiability of g at X for Y if and only if θ enjoys the same property at $\lambda(X)$ for $\lambda(Y)$. It is not clear yet to us whether such an equivalence can be achieved via our approach. Note that our main result in this section provides the equivalence for parabolic regularity of g and θ. It is worth mentioning here that parabolic regularity is strictly stronger than twice epi-differentiability and has no counterpart for the proximal mapping. Thus, Theorem 4 can't be derived from Ding et al. [11, theorem 3].

We close this section by establishing a characterization of twice semidifferentiability of the spectral function g in (1) when the symmetric function θ therein is convex. Recall from Rockafellar and Wets [25, exercise 13.7] that a function $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ is called *twice semidifferentiable* at $\bar{x}$ if there exists a continuous function h, which is positive homogeneous of degree 2, and

$$f(x) = f(\bar{x}) + \mathrm{d}f(\bar{x})(x - \bar{x}) + \frac{1}{2}h(x - \bar{x}) + o(\|x - \bar{x}\|^2).$$

In this case, h is called the *second semiderivative* of f at $\bar{x}$ and is denoted by $\mathrm{d}^2 f(\bar{x})$.

Corollary 5. *Assume that $X \in \mathbf{S}^n$ and $\mu_1 > \cdots > \mu_r$ are the distinct eigenvalues of X that $\theta : \mathbf{R}^n \to \mathbf{R}$ is a differentiable symmetric convex function. Then, θ is twice semidifferentiable at $\lambda(X)$ if and only if the spectral function $g = \theta \circ \lambda$ is twice semidifferentiable at X. Moreover, in this case, we have*

$$\mathrm{d}^2 g(X)(H) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)) + 2\sum_{m=1}^{r} \langle \Lambda(Y)_{\alpha_m\alpha_m}, U_{\alpha_m}^\top H(\mu_m I - X)^\dagger H U_{\alpha_m}\rangle,$$

where α_m, $m = 1, \ldots, r$, come from (7), $U \in \mathbf{O}^n(X) \cap \mathbf{O}^n(Y)$, $\Lambda(Y) = \operatorname{Diag}\,(\lambda(Y))$, and $H \in \mathbf{S}^n$.

Proof. Observe first that because θ is convex and finite-valued, it is locally Lipschitz continuous on $\mathbf{R}^n$. Moreover, because θ is differentiable, we deduce from Lewis [17, theorem 1.1] that g is differentiable at X. Suppose first that θ is twice semidifferentiable at $\lambda(X)$. Because θ is differentiable, it follows from twice semidifferentiability of θ that the second subderivative of θ coincides with its second semiderivative, namely,

$$\mathrm{d}^2\theta(\lambda(X), \nabla\theta(\lambda(X)))(\lambda'(X;H)) = \mathrm{d}^2\theta(\lambda(X))(\lambda'(X;H)) \quad \text{for all } H \in \mathbf{S}^n.$$

This, combined with the formula of the second subderivative of g in Theorem 4(a), tells us that for any $H \in \mathbf{S}^n$, $\mathrm{d}^2 g(X, \nabla g(X))(H)$ is always finite. By Mohammadi and Sarabi [20, example 4.7(d)], θ is parabolically epi-differentiable and parabolically regular at $\lambda(X)$ for $\nabla\theta(\lambda(X))$. Thus, it results from Corollary 4 that g is twice epi-differentiable at X for $\nabla g(X)$. Combining these with Rockafellar [24, theorem 4.3] tells us that g is twice semidifferentiable at X and that $\mathrm{d}^2 g(X, \nabla g(X))(H) = \mathrm{d}^2 g(X)(H)$ for any $H \in \mathbf{S}^n$. The latter, coupled with Theorem 4(a), proves the claimed formula for the second semiderivative of g at X. Conversely, assume that g is twice semidifferentiable at X. We know from (19) that the symmetric function θ can be represented as $\theta = g \circ F$ with $F(x) := \operatorname{Diag}\,(x)$ for all $x \in \mathbf{R}^n$. Because F is always twice differentiable and g is twice semidifferentiable at X, we conclude from Mohammadi et al. [22, proposition 8.2(i)] that θ is twice semidifferentiable at $\lambda(X)$, which completes the proof. □

One can also show similar to the proof of Corollary 5 that θ has a quadratic expansion at $\lambda(X)$ if and only if $\theta \circ \lambda$ enjoys the same property at X. Note that it was shown in Lewis and Sendov [19, theorem 3.3] (see also Drusvyatskiy and Paquette [13] for a simplified proof) that twice differentiability of g and θ are also equivalent. Whether such a result can be derived from our established theory in this section remains an open question for our future research.

6. Conclusion and Future Research

In this paper, we developed a second-order theory of generalized differentiation for spectral functions. Our results rely heavily upon the metric subregularity constraint qualification, which automatically holds for this setting. Our main focus was to characterize parabolic regularity of this class of functions when they are convex. Moreover, we were able to calculate their second subderivative.

Our results raise several questions for our future search. First and foremost is the extension of our established theory for subdifferentially regular spectral functions. This will allow us to provide a unified umbrella, under which all the available results for both convex and nonconvex spectral functions can be covered by our approach. Also, it is interesting to see whether a similar characterization of parabolic regularity can be achieved for twice epi-differentiability of spectral functions. Such a characterization can be obtained from Ding et al. [11, theorem 3] for convex spectral functions. However, we can't use Ding et al. [11] to obtain a similar characterization for nonconvex spectral functions. It is also important to see whether our results can be utilized to characterize twice differentiability of spectral functions, which was previously obtained by Lewis and Sendov [19, theorem 3.3].

References

[1] Beck A (2017) *First-Order Methods in Optimization* (Society for Industrial and Applied Mathematics, Philadelphia).
[2] Ben-Tal A, Zowe J (1982) A unified theory of first- and second-order conditions for extremum problems in topological vector spaces. *Math. Program.* 19:39–76.
[3] Bonnans JF, Shapiro A (2000) *Perturbation Analysis of Optimization Problems* (Springer, New York).
[4] Borwein JM, Lewis AS (2008) *Convex Analysis and Nonlinear Optimization*, 2nd ed. (Springer, New York).
[5] Chan ZX, Sun DF (2008) Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. *SIAM J. Optim.* 19:370–396.
[6] Cui Y, Ding C (2019) Nonsmooth composite matrix optimization: Strong regularity, constraint nondegeneracy and beyond. Preprint, submitted July 30, https://arxiv.org/abs/1907.13253.
[7] Cui Y, Ding C, Zhao XY (2019) Quadratic growth conditions for convex matrix optimization problems associated with spectral functions. *SIAM J. Optim.* 27:2332–2355.
[8] Daniilidis A, Lewis AS, Malick J, Sendov H (2008) Prox-regularity of spectral functions and spectral sets. *J. Convex Anal.* 15(3):547–560.
[9] Davis C (1957) All convex invariant functions of Hermitian matrices. *Arch. Math.* 8:276–278.
[10] Ding C (2012) An introduction to a class of matrix optimization problems. PhD thesis, National University of Singapore, Singapore.
[11] Ding C, Sun D, Sun J, Toh K (2018) Spectral operators of matrices. *Math. Program.* 168:509–531.
[12] Dontchev AL, Rockafellar RT (2014) *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd ed. (Springer, New York).
[13] Drusvyatskiy D, Paquette C (2018) Variational analysis of spectral functions simplified. *J. Convex Anal.* 25(1):119–134.
[14] Gfrerer H, Outrata JV (2016) On Lipschitzian properties of implicit multifunctions. *SIAM J. Optim.* 26:2160–2189.
[15] Gfrerer H, Ye JJ, Zhou J (2022) Second-order optimality conditions for non-convex set-constrained optimization problems. *Math. Oper. Res.* 47:2344–2365.
[16] Hang NTV, Sarabi ME (2021) Local convergence analysis of augmented Lagrangian methods for piecewise linear-quadratic composite optimization problems. *SIAM J. Optim.* 31(4):2665–2694.
[17] Lewis AS (1996) Convex analysis on the Hermitian matrices. *SIAM J. Optim.* 6:164–177.
[18] Lewis AS (1999) Nonsmooth analysis of eigenvalues. *Math. Program.* 84:1–24.
[19] Lewis AS, Sendov HS (2001) Twice differentiable spectral functions. *SIAM J. Matrix Anal. Appl.* 23:368–386.
[20] Mohammadi A, Sarabi ME (2020) Twice epi-differentiability of extended real-valued functions with applications in composite optimization. *SIAM J. Optim.* 30:2379–2409.
[21] Mohammadi A, Mordukhovich BS, Sarabi ME (2021) Parabolic regularity in geometric variational analysis. *Trans. Amer. Soc.* 374:1711–1763.
[22] Mohammadi A, Mordukhovich BS, Sarabi ME (2021) Variational analysis of composite models with applications to continuous optimization. *Math. Oper. Res.* 47:397–426.
[23] Rockafellar RT (1988) First- and second-order epi-differentiability in nonlinear programming. *Trans. Amer. Math. Soc.* 307:75–108.
[24] Rockafellar RT (2000) Second-order convex analysis. *J. Nonlinear Convex Anal.* 1:1–16.
[25] Rockafellar RT, Wets RJ-B (1998) *Variational Analysis* (Springer, Berlin).
[26] Torki M (1998) First- and second-order epi-differentiability in eigenvalue optimization. *J. Math. Anal. Appl.* 234:391–416.
[27] Torki M (2001) Second-order directional derivatives of all eigenvalues of a symmetric matrix. *Nonlinear Anal.* 46:1133–1150.
[28] Zhang L, Zhang N, Xiao X (2013) On the second-order directional derivatives of singular values of matrices and symmetric matrix-valued functions. *Set-Valued Variational Anal.* 21:557–586.