Math. Program., Ser. B 104, 263–292 (2005)

Digital Object Identifier (DOI) 10.1007/s10107-005-0616-1

James V. Burke · Adrian S. Lewis · Michael L. Overton

Variational analysis of functions of the roots of polynomials

Dedicated to R. Tyrrell Rockafellar on the occasion of his 70th birthday. Terry is one of those rare individuals who combine a broad vision, deep insight, and the outstanding writing and lecturing skills crucial for engaging others in his subject. With these qualities he has won universal respect as a founding father of our discipline. We, and the broader mathematical community, owe Terry a great deal. But most of all we are personally thankful to Terry for his friendship and guidance.

Received: May 2, 2004 / Accepted: February 25, 2005
Published online: July 14, 2005 –

Abstract. The Gauss-Lucas Theorem on the roots of polynomials nicely simplifies the computation of the subderivative and regular subdifferential of the abscissa mapping on polynomials (the maximum of the real parts of the roots). This paper extends this approach to more general functions of the roots. By combining the Gauss-Lucas methodology with an analysis of the splitting behavior of the roots, we obtain characterizations of the subderivative and regular subdifferential for these functions as well. In particular, we completely characterize the subderivative and regular subdifferential of the radius mapping (the maximum of the moduli of the roots). The abscissa and radius mappings are important for the study of continuous and discrete time linear dynamical systems.

1. Introduction

Let $\mathcal{P}^n$ denote the linear space of complex polynomials of degree n or less. Define the *root mapping* on $\mathcal{P}^n$ to be the multifunction $\mathcal{R} : \mathcal{P}^n \to \mathbb{C}$ given by

$$\mathcal{R}(p) = \{\lambda \mid p(\lambda) = 0\},$$

and let $\phi : \mathbb{C} \to I\!R \cup \{+\infty\} = \bar{I\!R}$ be a lower semi-continuous convex function. We are concerned with the variational properties of functions $\hat{\phi} : \mathcal{P}^n \to I\!R \cup \{\pm\infty\}$ defined as

$$\hat{\phi}(p) = \sup\{\phi(\lambda) \mid \lambda \in \mathcal{R}(p)\} . \tag{1}$$

The abscissa and radius mappings on $\mathcal{P}^n$ are obtained by taking $\phi(\lambda) = \operatorname{Re}\lambda$ and $\phi(\lambda) = |\lambda|$, respectively. Indeed, the abscissa and radius mappings are the primary motivation for

J.V. Burke: Department of Mathematics, University of Washington, Seattle, WA 98195, USA. e-mail: burke@math.washington.edu

Research supported in part by the National Science Foundation Grant DMS-0203175.

A. S. Lewis: School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853, USA. e-mail: aslewis@orie.cornell.edu

Research supported in part by the Natural Sciences and Engineering Research Council of Canada.

M.L. Overton: Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA. e-mail: overton@cs.nyu.edu

Research supported in part by the National Science Foundation Grant DMS-0412049.

Mathematics Subject Classification (1991): 90C46, 49K40, 65K05, 15A42

this study. The variational behavior of these functions is important for our understanding of the stability properties of continuous and discrete time dynamical systems [1, 2].

The functions defined by (1) possess a very rich variational structure. In addition, these functions are related to important applied problems. Consequently, they offer an ideal setting in which to test the utility and robustness of any theory for analyzing the variational structure of nonsmooth functions. We study the variational behavior of the class (1) on the set $\mathcal{M}^n$ of polynomials of degree n. This set is an open dense subset of the linear space $\mathcal{P}^n$ (endowed with the topology of pointwise convergence). Note that on the set $\mathcal{M}^n$ the *sup* in (1) can be replaced by *max*. The supremum in (1) is only required for constant polynomials. We also note that the non-constant members of the class (1) are never locally Lipschitz on $\mathcal{M}^n$ and are always unbounded in the neighborhood of any point on the boundary of $\mathcal{M}^n$. The non-Lipschitzian behavior is seen by considering the family of polynomials $p_\epsilon(\lambda) = (\lambda - \lambda_0)^n - \epsilon$. To see that $\hat{\phi}$ must be unbounded on the boundary of $\mathcal{M}^n$ recall that for any non-constant convex function ϕ there must exist a nonzero direction $a \in \mathbb{C}$ for which $\phi(\tau a) \to +\infty$ as $\tau \uparrow \infty$. If $p \in \mathcal{P}^n$ is any polynomial of degree less than n, the polynomial defined by $q_\epsilon(\lambda) = (1 - a^{-1}\epsilon\lambda)p(\lambda)$ is in $\mathcal{P}^n$, and satisfies $q_\epsilon \to p$ as $\epsilon \searrow 0$. Moreover, $\epsilon^{-1}a \in \mathcal{R}(q_\epsilon)$ for all $\epsilon > 0$. Therefore, $\hat{\phi}(q_\epsilon) \to +\infty$ as $\epsilon \searrow 0$. It is this essential unboundedness of the roots on the boundary of $\mathcal{M}^n$ that motivates the restriction to $\mathcal{M}^n$. On $\mathcal{M}^n$ the roots of polynomials are continuous functions of their coefficients, so the functions $\hat{\phi}$ defined in (1) are lower semi-continuous on $\mathcal{M}^n$.

We use the tools developed in [6, 7, 12, 14] to study the variational properties of $\hat{\phi}$. Our earlier work demonstrates that these techniques are well suited to applications in stability theory [1–5]. In addition, we make fundamental use of a classical result originally due to Gauss and commonly known as the Gauss-Lucas Theorem. This result establishes a beautiful and elementary convexity relationship between the roots of a polynomial and the roots of its derivative.

Theorem 1. [Gauss-Lucas] *All critical points of a non-constant polynomial p lie in the convex hull H of the set of roots of p. If the roots of p are not collinear, no critical point of p lies on the boundary of H unless it is a multiple root of p.*

The Gauss-Lucas Theorem implies the following chain of inclusions for any polynomial of degree n:

$$\operatorname{conv}\mathcal{R}(p^{(n-1)}) \subset \operatorname{conv}\mathcal{R}(p^{(n-2)}) \subset \cdots \subset \operatorname{conv}\mathcal{R}(p') \subset \operatorname{conv}\mathcal{R}(p), \tag{2}$$

where conv S denotes the convex hull of the set S.

Theorem 1 is referred to by Gauss as early as 1830 and has been rediscovered many times. In 1879, Lucas [9, 10] published a refinement of Gauss's result. For more on the Gauss-Lucas Theorem and its uses see Marden [11].

With regard to the chain (2), it is useful to observe that

$$\hat{\phi}(p) = \sup\{\phi(\lambda) \mid \lambda \in \mathcal{R}(p)\} = \sup\{\phi(\lambda) \mid \lambda \in \operatorname{conv}\mathcal{R}(p)\}\ . \tag{3}$$

To see this note that the second supremum is clearly greater than or equal to the first so we need only show the reverse inequality. Let $\lambda \in \operatorname{conv}\mathcal{R}(p)$, then, by Caratheodory, there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathcal{R}(p)$ and $0 \leq \mu_1, \mu_2, \mu_3$ with $\mu_1 + \mu_2 + \mu_3 = 1$ such that $\lambda = \mu_1\lambda_1 + \mu_2\lambda_2 + \mu_3\lambda_3$. But then by convexity

$$\phi(\lambda) \leq \mu_1\phi(\lambda_1) + \mu_2\phi(\lambda_2) + \mu_3\phi(\lambda_3) \leq \max_{j=1,2,3} \phi(\lambda_j),$$

so that the second supremum in (3) is less than or equal to the first.

In [5], Burke and Overton investigate the variational properties of the abscissa mapping using an approach modeled on work of Levantovskii [8]. However, this approach is difficult and lengthy, and provides little insight into the underlying variational geometry. Furthermore, extending this approach to other functions of the roots of polynomials would be a daunting task at best. In [3], an approach based on the Gauss-Lucas Theorem is introduced to simplify the derivation of the tangent cone to the epigraph of the abscissa mapping at the polynomial $p(\lambda) = (\lambda - \lambda_0)^n$. This derivation is one of the two most difficult technical hurdles in [5]. The second is the verification of subdifferential regularity.

In this paper, we apply the Gauss-Lucas ideas in [3] to the class of functions given by (1). As in [3], we focus on the computation of the tangent cone to the epigraph at the polynomials

$$e_{(n,\lambda_0)}(\lambda) = (\lambda - \lambda_0)^n.$$

We recover our earlier result for the abscissa mapping and obtain the corresponding result for the radius mapping which is stated below. Here and throughout, we denote the complex conjugate of the complex scalar z by $\bar{z}$.

Theorem 2. *Let* $\mathsf{r} : \mathcal{P}^n \to I\!R$ *denote the radius mapping on* $\mathcal{P}^n$*:*

$$\mathsf{r}(p) = \max\{|\lambda| \mid \lambda \in \mathcal{R}(p)\}.$$

Let $(v, \eta) \in \mathcal{P}^n \times I\!R$ *be such that* $v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}$.

(i) (v, η) *is an element of the tangent cone to the epigraph of* r *at the polynomial* $p(\lambda) = \lambda^n$ *if and only if*

$$\eta \geq \frac{1}{n}|b_1| \quad \textit{and} \quad 0 = b_k,\ k = 2, 3, \ldots, n.$$

(ii) (v, η) *is an element of the tangent cone to the epigraph of* r *at the polynomial* $p(\lambda) = (\lambda - \lambda_0)^n$ *with* $\lambda_0 \neq 0$ *if and only if*

$$\eta \geq \frac{1}{n|\lambda_0|}\left[|b_2| - \operatorname{Re}\bar{\lambda}_0 b_1\right],$$
$$0 = \operatorname{Re}\bar{\lambda}_0\sqrt{-b_2}, \quad \textit{and} \quad 0 = b_k,\ k = 3, \ldots, n.$$

The notation and definitions follow those established in [14]. The definitions of the terms *epigraph* and *tangent cone* used in Theorem 2 appear in the next section. Notation specific to the study of polynomials on the complex plane is introduced below.

Recall that the complex plane $\mathbb{C}$ is a Euclidean space when endowed with the usual real inner product $\langle \omega \, , \, \lambda \rangle = \mathrm{Re}\, \bar{\omega}\lambda$. By extension, $\mathbb{C}^n$ is also a Euclidean space when given the real inner product

$$\left\langle (\omega_1, \ldots, \omega_n)^T \, , \, (\lambda_1, \ldots, \lambda_n)^T \right\rangle = \sum_{k=1}^{n} \langle \omega_k \, , \, \lambda_k \rangle \, .$$

Given $\lambda_0 \in \mathbb{C}$ we define the basis $\left\{ e_{(k,\lambda_0)} \mid k = 0, 1, \ldots, n \right\}$ for $\mathcal{P}^n$, where

$$e_{(k,\lambda_0)}(\lambda) = (\lambda - \lambda_0)^k, \; k = 0, 1, \ldots, n.$$

Each such basis defines a real inner product (or *duality pairing*) on $\mathcal{P}^n$:

$$\langle p \, , \, q \rangle_{\lambda_0} = \mathrm{Re} \sum_{k=0}^{n} \bar{a}_k b_k,$$

where $p = \sum_{k=0}^{n} a_k e_{(k,\lambda_0)}$ and $q = \sum_{k=0}^{n} b_k e_{(k,\lambda_0)}$. In the case $n = 0$, we recover the real inner product on $\mathbb{C}$. When $\lambda_0 = 0$, we drop the subscript on the inner product and simply write $\langle \cdot \, , \, \cdot \rangle$. Note that this family of inner products behaves continuously in $p, \; q$, and λ_0 in the sense that the mapping $(p, q, \lambda_0) \to \langle p \, , \, q \rangle_{\lambda_0}$ is continuous on $\mathcal{P}^n \times \mathcal{P}^n \times \mathbb{C}$. To see this simply note that

$$\langle p \, , \, q \rangle_{\lambda} = \mathrm{Re} \left(\sum_{k=0}^{n} \overline{\frac{p^{(k)}(\lambda)}{k!}} \, \frac{q^{(k)}(\lambda)}{k!} \right).$$

The inner product on a Euclidean space gives rise to a norm $|\cdot|$ in the usual way by setting $|x| = \sqrt{\langle x \, , \, x \rangle}$.

Given a mapping $\phi \; : \; \mathbb{C} \to \bar{I\!R}$, we define the mapping $\tilde{\phi} \; : \; I\!R^2 \to \bar{I\!R}$ by the composition $\tilde{\phi} = \phi \circ \Theta$, where $\Theta \; : \; I\!R^2 \to \mathbb{C}$ is the linear transformation

$$\Theta(x) = x_1 + \mathrm{i}\, x_2, \tag{4}$$

$\mathrm{i} \in \mathbb{C}$ denoting the imaginary unit. If we endow $I\!R^2$ with its usual inner product and $\mathbb{C}$ with the inner product above, we have

$$\Theta^{-1}\mu = \Theta^{*}\mu = \begin{bmatrix} \mathrm{Re} \;\, \mu \\ \mathrm{Im} \;\, \mu \end{bmatrix}.$$

We say that ϕ is *differentiable in the real sense* if $\tilde{\phi}$ is differentiable, in which case the derivative of ϕ is given by the chain rule as

$$\phi'(\zeta) = \Theta \nabla \tilde{\phi}(\Theta^{*}\zeta).$$

Here $\nabla\tilde{\phi}$ denotes the gradient of $\tilde{\phi}$. Similarly, we say that ϕ is twice differentiable in the real sense if $\tilde{\phi}$ is twice differentiable and again the chain rule gives

$$\phi''(\zeta)\delta = \Theta \nabla^2 \tilde{\phi}(\Theta^{*}\zeta)\Theta^{*}\delta, \tag{5}$$

where $\nabla^2\tilde{\phi}$ denotes the Hessian of $\tilde{\phi}$. Since these are the only notions of differentiability we employ, we omit the qualifying phrase *in the real sense* when specifying that a function from $\mathbb{C}$ to $\bar{I\!R}$ is differentiable or twice differentiable. We also make use of the following notation:

$$\phi'(\zeta; \delta) = \left\langle \phi'(\zeta) \, , \, \delta \right\rangle \quad \text{and} \quad \phi''(\zeta; \omega, \delta) = \left\langle \omega \, , \, \phi''(\zeta)\delta \right\rangle.$$

2. The tangent cone and the subderivative

We use the tools in [7, 14] to describe the variational geometry of the function $\hat{\phi}$. Recall that the epigraph of a function f mapping a Euclidean space E into the extended real numbers $\bar{I\!R} = I\!R \cup \{+\infty\}$ is the subset of $E \times I\!R$ given by

$$\text{epi}\,(f) = \{(x, \mu) \mid f(x) \le \mu\}\,, \tag{6}$$

and the domain is the set

$$\text{dom}\,(f) = \{x \mid f(x) < \infty\}\,.$$

The fundamental variational object for f is the tangent cone to its epigraph [14, Definition 6.1]. Given a set C in a linear space S and a point $x \in C$, the tangent cone to C at x is defined to be the set

$$T_C(x) = \left\{ d \;\middle|\; \begin{array}{c} \exists\{x^k\} \subset C,\ \{t_k\} \subset I\!R \text{ such that} \\ x^k \to x,\ t_k \downarrow 0,\ \text{and } t_k^{-1}(x^k - x) \to d \end{array} \right\}.$$

The tangent cone to epi (f) at a point can be viewed as the epigraph of another function called the subderivative of f at x. It is denoted by $df(x)$ [14, Theorem 8.2]:

$$\text{epi}\,(df(x)) = \{(w, \eta) \mid df(x)(w) \le \eta\} = T_{\text{epi}\,(f)}(x, f(x)) \tag{7}$$

for all $x \in \text{dom}\,(f) = \{x \mid f(x) < +\infty\}$. The subderivative generalizes the notion of a directional derivative as seen by the following alternative formula [14, Definition 8.1]:

$$df(x)(w) = \liminf_{\substack{t \searrow 0 \\ w' \to w}} \frac{f(x + tw') - f(x)}{t}\,. \tag{8}$$

In the case of the function class (1), the computation of the tangent cone is simplified due to the fact that $(p, \eta) \in \text{epi}\,(\hat{\phi})$ if and only if $(\zeta p, \eta) \in \text{epi}\,(\hat{\phi})$ for every nonzero complex scalar ζ.

Lemma 1. *Define $\mathcal{M}_1^n \subset \mathcal{M}^n$ to be the set of monic polynomials of degree n:*

$$\mathcal{M}_1^n = \left\{ e_{(n,0)} + q \;\middle|\; q \in \mathcal{P}^{n-1} \right\}.$$

Define $\hat{\phi}_1 : \mathcal{P}^n \to \bar{I\!R}$ by

$$\hat{\phi}_1(p) = \begin{cases} \hat{\phi}(p)\ , \text{ if } p \in \mathcal{M}_1^n, \\ +\infty\ \ , \text{ otherwise.} \end{cases}$$

Let $\lambda_0 \in \text{dom}\,(\phi)$, $b_k \in \mathbb{C},\ k = 0, 1, \ldots, n$, $\eta \in I\!R$, and set

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)} \quad \text{and} \quad \tilde{v} = \sum_{k=1}^{n} b_k e_{(n-k,\lambda_0)}.$$

Then

$$(v, \eta) \in T_{\text{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$$

if and only if

$$(\tilde{v}, \eta) \in T_{\text{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right).$$

Remark 1. The lemma shows that

$$T_{\text{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right) = \mathbb{C}\, e_{(n,\lambda_0)} + T_{\text{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right),$$

where $\mathbb{C}\, e_{(n,\lambda_0)} = \left\{\xi e_{(n,\lambda_0)} \mid \xi \in \mathbb{C}\right\}$. Therefore, we can restrict our analysis of the tangent cone to sequences that lie in the set $\mathcal{M}_1^n$. This is the approach taken in [4]. Here we work on the seemingly more general space $\mathcal{M}^n$ in order to simplify applications.

Proof. Suppose $(v, \eta) \in T_{\text{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$, that is, there exists a sequence $\xi_j \downarrow 0$ such that

$$(e_{(n,\lambda_0)} + \xi_j v + o(\xi_j), \phi(\lambda_0) + \xi_j \eta + o(\xi_j)) \in \text{epi}\,(\hat{\phi}),$$

or equivalently,

$$(e_{(n,\lambda_0)} + \frac{\xi_j \tilde{v}}{1 + \xi_j b_0} + o(\xi_j), \phi(\lambda_0) + \xi_j \eta + o(\xi_j)) \in \text{epi}\,(\hat{\phi}_1).$$

Hence $(\tilde{v}, \eta) \in T_{\text{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$.

Conversely, suppose $(\tilde{v}, \eta) \in T_{\text{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$. By definition, there exists $\xi_j \downarrow 0$ such that

$$(e_{(n,\lambda_0)} + \xi_j \tilde{v} + o(\xi_j), \phi(\lambda_0) + \xi_j \eta + o(\xi_j)) \in \text{epi}\,(\hat{\phi}_1). \tag{9}$$

Multiplying $e_{(n,\lambda_0)} + \xi_j \tilde{v} + o(\xi_j)$ by $(1 + \xi_j b_0)$ gives

$$(1 + \xi_j b_0)(e_{(n,\lambda_0)} + \xi_j \tilde{v} + o(\xi_j)) = e_{(n,\lambda_0)} + \xi_j v + o(\xi_j).$$

That is, (9) is equivalent to

$$(e_{(n,\lambda_0)} + \xi_j v + o(\xi_j), \phi(\lambda_0) + \xi_j \eta + o(\xi_j)) \in \text{epi}\,(\hat{\phi})\,.$$

Therefore, $(v, \eta) \in T_{\text{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$ which proves the result. □

Next recall that the epigraph of the convex function ϕ as well as all of the lower level sets

$$\text{lev}_\phi(\mu) = \{\lambda \mid \phi(\lambda) \le \mu\}$$

are convex sets [13]. This allows us to apply the Gauss-Lucas Theorem in a powerful way. Since

$$\mu \ge \hat{\phi}(p) \iff \mathcal{R}(p) \subset \text{lev}_\phi(\mu), \tag{10}$$

the Gauss-Lucas Theorem yields the following chain of inclusions for any polynomial of degree n providing $(p, \mu) \in \text{epi}\,(\hat{\phi})$:

$$\text{conv}\,\mathcal{R}(p^{(n-1)}) \subset \cdots \subset \text{conv}\,\mathcal{R}(p') \subset \text{conv}\,\mathcal{R}(p) \subset \text{lev}_\phi(\mu). \tag{11}$$

This system of inclusions along with subdifferential information about the function ϕ provide the basis for a set of necessary conditions for a pair $(v, \eta) \in \mathcal{P}^n \times I\!R$ to be an element of the tangent cone to epi $(\hat{\phi})$.

The convexity of ϕ implies that ϕ is directionally differentiable in all directions at every point in $\lambda_0 \in \operatorname{dom}(\phi)$ [13] and one has

$$\phi'(\lambda_0; \xi) = \lim_{\tau \downarrow 0} \frac{\phi(\lambda_0 + \tau\xi) - \phi(\lambda_0)}{\tau} = \inf_{\tau > 0} \frac{\phi(\lambda_0 + \tau\xi) - \phi(\lambda_0)}{\tau}.$$

Taking $\tau = 1$ and $\xi = \lambda - \lambda_0$ in the right hand side of this expression gives the subdifferential inequality

$$\phi(\lambda) \geq \phi(\lambda_0) + \phi'(\lambda_0; \lambda - \lambda_0).$$

A vector ω is a subgradient of ϕ, written $\omega \in \partial\phi(\lambda)$, if and only if

$$\phi(\lambda) \geq \phi(\lambda_0) + \langle \omega\,,\ \lambda - \lambda_0 \rangle\ \forall \lambda \in \mathbb{C}\,. \tag{12}$$

The subdifferential $\partial\phi(\lambda_0)$ is always a closed convex set, although it may be empty at points on the boundary of the set $\operatorname{dom}(\phi)$. The subdifferential is related to the directional derivative of ϕ by the formula

$$\phi'(\lambda_0; \lambda) = \sup\left\{\langle z\,,\ \lambda \rangle \mid z \in \partial\phi(\lambda_0)\right\}\ . \tag{13}$$

Therefore, at points

$$\lambda_0 \in \operatorname{dom}(\partial\phi) = \{\lambda \mid \partial\phi(\lambda) \neq \emptyset\}\,,$$

we have $\phi'(\lambda_0; \cdot) : \mathbb{C} \to I\!R \cup \{+\infty\}$ is a lower semi-continuous convex function. Since $\phi'(\lambda_0; \cdot)$ is a convex function, we can use (1) to define

$$\hat{\phi}_{\lambda_0} = \widehat{\phi'(\lambda_0; \cdot)}.$$

That is, we define the function $\hat{\phi}_{\lambda_0} : \mathcal{P}^n \to I\!R \cup \{\pm\infty\}$ by

$$\hat{\phi}_{\lambda_0}(q) = \sup\left\{\phi'(\lambda_0; \lambda) \mid q(\lambda) = 0\right\},$$

and observe that by (13), (3), and (11),

$$\begin{aligned}\hat{\phi}_{\lambda_0}(q) &= \sup\left\{\langle z\,,\ \lambda)\rangle \mid (z, \lambda) \in \partial\phi(\lambda_0) \times \operatorname{conv}\mathcal{R}(q)\right\} \\ &\geq \sup\left\{\langle z\,,\ \lambda)\rangle \;\middle|\; (z, \lambda) \in \partial\phi(\lambda_0) \times \operatorname{conv}\mathcal{R}(q')\right\} \\ &= \hat{\phi}_{\lambda_0}(q').\end{aligned} \tag{14}$$

The next lemma shows how to extend the subdifferential inequality for ϕ to the function $\hat{\phi}$ by using the function $\hat{\phi}_{\lambda_0}$.

Lemma 2. *Given* $0 < \tau \in I\!R$ *and* $\lambda_0 \in \operatorname{dom}(\phi)$*, define the linear transformation* $\mathcal{S}_{[\tau,\lambda_0]} : \mathcal{P}^n \to \mathcal{P}^n$ *by* $\mathcal{S}_{[\tau,\lambda_0]}(p)(\lambda) = p(\lambda_0 + \tau\lambda)$*. Then, for every* $p \in \mathcal{P}^n$*,*

$$\hat{\phi}(p) \geq \phi(\lambda_0) + \tau\hat{\phi}_{\lambda_0}([\mathcal{S}_{[\tau,\lambda_0]}(p)]^{(\ell)}), \tag{15}$$

for $\ell = 0, 1, \ldots, (\deg(p) - 1)$.

Proof. For the case $\ell = 0$, we have

$$\begin{aligned}
\hat{\phi}(p) &= \max\{\phi(\lambda) \mid p(\lambda) = 0\} \\
&= \max\left\{\phi(\lambda) \;\middle|\; \mathcal{S}_{[\tau,\lambda_0]}(p)(\gamma) = 0,\ \lambda = \lambda_0 + \tau\gamma\right\} \\
&= \max\left\{\phi(\lambda_0 + \tau\gamma) \;\middle|\; \mathcal{S}_{[\tau,\lambda_0]}(p)(\gamma) = 0\right\} \\
&\geq \max\left\{\phi(\lambda_0) + \tau\phi'(\lambda_0;\gamma) \;\middle|\; \mathcal{S}_{[\tau,\lambda_0]}(p)(\gamma) = 0\right\} \\
&= \phi(\lambda_0) + \tau\hat{\phi}_{\lambda_0}(\mathcal{S}_{[\tau,\lambda_0]}(p)).
\end{aligned}$$

The remaining cases follow immediately from (14) and (11). □

We now translate the content of Lemma 2 into statements about the coefficients of the underlying polynomials. Consider the polynomial

$$p = \sum_{k=0}^{n} a_k e_{(n-k,\lambda_0)}, \tag{16}$$

with $a_0 \neq 0$. The ℓth derivative of this polynomial is given by

$$p^{(\ell)} = \ell! \sum_{k=0}^{n-\ell} b(n-k,\ell) a_k e_{(n-(k+\ell),\lambda_0)}, \tag{17}$$

where $b(n,k)$ with $k \leq n$ are the binomial coefficients $b(n,k) = \frac{n!}{k!(n-k)!}$. Applying the operator $\mathcal{S}_{[\tau,\lambda_0]}$ to p yields

$$\mathcal{S}_{[\tau,\lambda_0]}(p) = \sum_{k=0}^{n} a_k \tau^{(n-k)} e_{(n-k,0)}, \tag{18}$$

and

$$[\mathcal{S}_{[\tau,\lambda_0]}(p)]^{(\ell)} = \ell! \tau^n \sum_{k=0}^{n-\ell} b(n-k,\ell) \tau^{-k} a_k e_{(n-(k+\ell),0)}, \tag{19}$$

for $\ell = 0, 1, 2, \ldots, (n-1)$ (the case $\ell = 0$ is just (18). With this notation, we have the following simple consequence of Lemma 2.

Lemma 3. *Let $p \in \mathcal{P}^n$ be as in (16) and let $t \in I\!R$ be positive. Then*

$$\hat{\phi}(p) \geq \phi(\lambda_0) + t^{1/s} \hat{\phi}_{\lambda_0}\left(\sum_{k=0}^{s} b(n-k,n-s) t^{-k/s} a_k e_{(s-k,0)}\right), \tag{20}$$

for $s = 1, \ldots, n$.

Proof. In (19) take $\ell = n - s$ and $\tau = t^{1/s}$ for $\ell = 0, 1, 2, \ldots, (n-1)$, or equivalently, $s = 1, \ldots, n$, to obtain

$$\mathcal{S}_{[\tau,\lambda_0]}(p)^{(n-s)} = \ell! t^{n/s} \sum_{k=0}^{s} b(n-k,n-s) t^{-k/s} a_k e_{(s-k,0)},$$

for $s = 1, \ldots, n$. Plugging this expression into (15) yields the result since $\hat{\phi}_{\lambda_0}(p) = \hat{\phi}_{\lambda_0}(\zeta p)$ for every nonzero complex number ζ. □

The main result of the section now follows.

Theorem 3. *Let* $\lambda_0 \in \operatorname{dom}(\partial\phi)$ *with* $\partial\phi(\lambda_0) \neq \{0\}$. *If*

$$(v, \eta) \in T_{\operatorname{epi}(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$$

with

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}, \tag{21}$$

then

$$\eta \geq \phi'(\lambda_0; -b_1/n), \tag{22}$$
$$0 = \left\langle g\,,\, \sqrt{-b_2}\right\rangle\ \forall\, g \in \partial\phi(\lambda_0), \quad \text{and} \tag{23}$$
$$0 = b_k,\ k = 3, \ldots, n. \tag{24}$$

Thus, in particular, if $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$, *then* $b_2 = 0$, *where for any* $D \subset \mathbb{C}$ *the set*

$$\operatorname{rspan}(D) = \left\{ \sum_{k=1}^{N} \tau_k x_k \mid N \in I\!N,\ \tau_k \in I\!R,\ x_k \in D\ k = 1, \ldots, N \right\}$$

is the real linear span of D.

Proof. Let $(v, \eta) \in T_{\operatorname{epi}(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$ with v given by (21).Then, by Lemma 1, $(\tilde{v}, \eta) \in T_{\operatorname{epi}(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$, where

$$\tilde{v} = \sum_{k=1}^{n} b_k e_{(n-k,\lambda_0)}.$$

Hence there exist sequences $t_j \downarrow 0$ and $\{(p_j, \mu_j)\} \in \operatorname{epi}(\hat{\phi}_1)$ such that

$$t_j^{-1}((p_j, \mu_j) - (e_{(n,\lambda_0)}, \phi(\lambda_0))) \to (\tilde{v}, \eta).$$

That is, there exists $\{(a_0^j, a_1^j, \ldots, a_n^j)\} \in \mathbb{C}^{n+1}$ such that

$$p_j = \sum_{k=0}^{n} a_k^j e_{(n-k,\lambda_0)},$$

$t_j^{-1}(\mu_j - \phi(\lambda_0)) \to \eta$, $a_0^j = 1$ for all $j = 1, 2, \ldots$, and $a_k^j \to 0$ with $t_j^{-1} a_k^j \to b_k$ for $k = 1, \ldots, n$. By applying Lemma 3 to p_j for each $j = 1, 2, \ldots$ with $t = t_j$ we obtain for $s = 1, 2, \ldots, n$

$$\mu_j \geq \phi(\lambda_0) + t_j^{1/s} \hat{\phi}_{\lambda_0} \left(\sum_{k=0}^{s} b(n-k, n-s) t_j^{-k/s} a_k^j e_{(s-k,0)} \right),$$

or equivalently,

$$t_j^{-1/s}(\mu_j - \phi(\lambda_0)) \geq \hat{\phi}_{\lambda_0}\left(\sum_{k=0}^{s} b(n-k, n-s) t_j^{-k/s} a_k^j e_{(s-k,0)}\right) \tag{25}$$

for each $s = 1, \ldots, n$. We now consider the limit as $j \to \infty$ in each of these inequalities. First observe that

$$\left[\sum_{k=0}^{s} b(n-k, n-s) t_j^{-k/s} a_k^j e_{(s-k,0)}\right] \to \left[b(n, n-s) e_{(s,0)} + b_s\right]$$

for $s = 1, \ldots, n$. Hence

$$\mathcal{R}\left(\sum_{k=0}^{s} b(n-k, n-s) t_j^{-k/s} a_k^j e_{(s-k,0)}\right) \to \mathcal{R}\left(b(n, n-s) e_{(s,0)} + b_s\right)$$

for $s = 1, \ldots, n$, since $a_0^j = 1$ for all $j = 1, 2, \ldots$ and the roots of a polynomial are a continuous function of its coefficients on $\mathcal{M}^n$. Therefore, the lower semi-continuity of $\hat{\phi}_{\lambda_0}$ and the inequalities (25) imply that

$$\eta \geq \hat{\phi}_{\lambda_0}(b(n, n-1) e_{(1,0)} + b_1) = \phi'(\lambda_0; -b_1/n) \tag{26}$$

which proves (22). For $s = 2, \ldots, n$, the inequalities (25) imply that

$$\begin{aligned} 0 &\geq \hat{\phi}_{\lambda_0}(b(n, n-s) e_{(s,0)} + b_s) \\ &= \max\left\{\phi'\left(\lambda_0; \left(\frac{-b_s}{b(n, n-s)}\right)^{1/s} \omega\right) \;\middle|\; \begin{array}{c} \omega = e^{2\pi k \mathrm{i}/s} \\ k = 0, \ldots, s-1 \end{array}\right\}. \end{aligned} \tag{27}$$

By assumption there exists $g \in \partial\phi(\lambda_0)$ with $g \neq 0$. Inequality (27) implies that

$$0 \geq \left\langle g, \left(\frac{-b_s}{b(n, n-s)}\right)^{1/s} \omega \right\rangle,$$

for $\omega = e^{2\pi k \mathrm{i}/s}$ $(k = 0, 1, \ldots, s-1)$. For $s = 3, \ldots, n$ this can only occur if $b_s = 0$ which gives (24). For $s = 2$ we have

$$0 \geq \left\langle g, \pm\left(\frac{-b_2}{b(n, n-2)}\right)^{1/2} \right\rangle \; \forall\, g \in \partial\phi(\lambda_0),$$

or equivalently, condition (23) holds. □

The conditions (23)and (24) play a key role in our subsequent analysis. Condition (24) is transparent, but condition (23) is not since it is nonlinear in b_2. In the following lemma we describe the underlying convexity of this condition. For this we use the notion of the *cone* generated by a subset of the complex plane. Given a set D contained in $\mathbb{C}$ we define the cone generated by D to be the set

$$\operatorname{cone}(D) = \{\tau x \mid \tau \in I\!R_+ \text{ and } x \in D\}.$$

If D is a convex subset of $\mathbb{C}$, then cone (D) is convex and rspan (D) = cone $(D) -$ cone (D). The *polar* of any convex cone K in $\mathbb{C}$ is the set

$$K^\circ = \{z \mid \langle z\,,\, x\rangle \le 0\; \forall\, x \in K\}\,,$$

and we have the relationship $\mathrm{cl}\,(K) = K^{\circ\circ}$.

Lemma 4. *Let* $\lambda_0 \in \mathrm{dom}\,(\partial\phi)$ *with* $\partial\phi(\lambda_0) \neq \{0\}$, *and let* $b_2 \in \mathbb{C}$. *Then the following statements are equivalent.*

(i) $\left\langle g\,,\, \sqrt{-b_2}\right\rangle = 0$ *for all* $g \in \partial\phi(\lambda_0)$.
(ii) Either $b_2 = 0$ *or* $\mathrm{rspan}\,(\sqrt{b_2}) = \mathrm{rspan}\,(\partial\phi(\lambda_0))$.
(iii) Either $b_2 = 0$ *or* $\mathrm{cone}\,(\partial\phi(\lambda_0)^2) = \mathrm{cone}\,(\{b_2\})$, *where*

$$\partial\phi(\lambda_0)^2 = \left\{g^2 \mid g \in \partial\phi(\lambda_0)\right\}.$$

In addition, the set of all $b_2 \in \mathbb{C}$ *satisfying (i) is given by the convex cone* $\mathcal{K}^\circ_{\lambda_0}$, *where*

$$\mathcal{K}_{\lambda_0} = -\mathrm{cone}\,(\partial\phi(\lambda_0)^2) + \mathrm{i}\,[\mathrm{rspan}\,(\partial\phi(\lambda_0)^2)] \tag{28}$$

so that

$$\mathcal{K}^\circ_{\lambda_0} = \begin{cases} \{0\} & ,\ \textit{if}\ \mathrm{rspan}\,(\partial\phi(\lambda_0)) = \mathbb{C},\ \textit{and} \\ \mathrm{cone}\,(\partial\phi(\lambda_0)^2) & ,\ \textit{otherwise.} \end{cases}$$

Proof. Observe that if $b_2 \neq 0$, then the condition in (i) implies that the real linear span rspan $(\partial\phi(\lambda_0))$ must be a line through the origin in $\mathbb{C}$ since $\partial\phi(\lambda_0) \neq \{0\}$. Note that any line through the origin can be written as rspan (v) for some $v \in \mathbb{C}$ and that the line perpendicular to this line in the real inner product is the line $\mathrm{i}\,[\mathrm{rspan}\,(v)]$. With these observations, we have the following string of equivalences:

$$\begin{gathered}
\left\langle g\,,\, \sqrt{-b_2}\right\rangle = 0\; \forall\, g \in \partial\phi(\lambda_0) \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or } \left\langle v\,,\, \mathrm{i}\sqrt{b_2}\right\rangle = 0\; \forall\, v \in \mathrm{rspan}\,(\partial\phi(\lambda_0)) \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or} \\
\langle v\,,\, w\rangle = 0\; \forall (v, w) \in \mathrm{rspan}\,(\partial\phi(\lambda_0)) \times (\,\mathrm{i}\,[\mathrm{rspan}\,(\sqrt{b_2})]) \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or } \mathrm{rspan}\,(\partial\phi(\lambda_0)) = \mathrm{rspan}\,(\sqrt{b_2}),
\end{gathered}$$

where the final equivalence follows since the lines $\mathrm{i}\,[\mathrm{rspan}\,(\partial\phi(\lambda_0))]$ and the line rspan $(\partial\phi(\lambda_0))$ are perpendicular whenever $b_2 \neq 0$. Therefore, (i) and (ii) are equivalent.

The equivalence of (ii) and (iii) follows from the following string of equivalences:

$$\begin{gathered}
\text{either } b_2 = 0 \text{ or } \operatorname{rspan}(\partial\phi(\lambda_0)) = \operatorname{rspan}(\sqrt{b_2}) \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or} \\
\forall\, g \in \partial\phi(\lambda_0)\ \exists\, \tau \in I\!R \text{ such that } g = \tau\sqrt{b_2} \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or} \\
\forall\, g \in \partial\phi(\lambda_0)\ \exists\, \tau \in I\!R \text{ such that } g^2 = \tau^2 b_2 \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or} \\
\operatorname{cone}(\{g^2\}) = \operatorname{cone}(\{b_2\}) \quad \forall\, g \in \partial\phi(\lambda_0) \setminus \{0\} \\
\Longleftrightarrow \\
\text{either } b_2 = 0 \text{ or } \operatorname{cone}(\partial\phi(\lambda_0)^2) = \operatorname{cone}(\{b_2\}).
\end{gathered}$$

Next note that the condition, $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$ is equivalent to the condition $\operatorname{rspan}(\partial\phi(\lambda_0)^2) = \mathbb{C}$ since the subdifferential $\partial\phi(\lambda_0)$ is a convex set. In this case the $\mathcal{K}_{\lambda_0} = \mathbb{C}$, and so $\mathcal{K}^\circ_{\lambda_0} = \{0\}$. Therefore, if $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$, then the set of all b_2 satisfying (23) equals $\mathcal{K}^\circ_{\lambda_0} = \{0\}$.

If $\operatorname{rspan}(\partial\phi(\lambda_0)) \neq \mathbb{C}$, the set $\operatorname{rspan}(\partial\phi(\lambda_0))$ is a line through the origin since $\partial\phi(\lambda_0) \neq \{0\}$. This line must equal the real linear span of any nonzero element of $\partial\phi(\lambda_0)$. Hence the set $\operatorname{cone}(\partial\phi(\lambda_0)^2)$ is a ray emanating from the origin (not a line), and the set $\mathcal{K}_{\lambda_0}$ is a convex cone in $\mathbb{C}$. An easy computation shows that the polar of this cone is the ray $\operatorname{cone}(\partial\phi(\lambda_0)^2)$. Hence, by (iii), the set of all b_2 satisfying (23) is contained in $\mathcal{K}^\circ_{\lambda_0}$. On the other hand, suppose $b_2 \in \mathcal{K}^\circ_{\lambda_0} = \operatorname{cone}(\partial\phi(\lambda_0)^2)$. Since $\operatorname{cone}(\partial\phi(\lambda_0)^2)$ is a ray emanating from the origin, we have

$$\mathcal{K}^\circ_{\lambda_0} = \operatorname{cone}(\partial\phi(\lambda_0)^2) = \operatorname{cone}(\{g^2\}) \quad \forall\, g \in \partial\phi(\lambda_0) \setminus \{0\}.$$

Hence, for each $g \in \partial\phi(\lambda_0) \setminus \{0\}$ there is a $\tau_g \geq 0$ such that $b_2 = \tau_g g^2$, or equivalently, $\sqrt{b_2} = \pm\sqrt{\tau_g}\, g$. Consequently,

$$\left\langle g\,,\ \sqrt{-b_2}\right\rangle = \pm\left\langle g\,,\ \mathrm{i}\sqrt{\tau_g}\, g\right\rangle = 0 \quad \forall\, g \in \partial\phi(\lambda_0) \setminus \{0\},$$

which shows that every $b_2 \in \mathcal{K}^\circ_{\lambda_0}$ satisfies (23) completing the proof that $\mathcal{K}^\circ_{\lambda_0}$ is precisely the set of all complex numbers that satisfy (23). $\square$

If ϕ is twice continuously differentiable with $\phi''(\lambda_0; \cdot, \cdot)$ positive definite, then Theorem 3 can be sharpened. For this we make use of the following technical result.

Lemma 5. *If H is a 2-by-2 real symmetric matrix with nonnegative trace then the function*

$$f(w) = \left\langle \Theta^{-1}\sqrt{w}\,,\ H\Theta^{-1}\sqrt{w}\right\rangle$$

is sublinear, i.e. positive homogeneous and subadditive (see (4) for the definition of Θ).

Proof. If

$$H = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

and $z = x + \mathrm{i}\, y$ with x and y real, then

$$\begin{aligned} f(z^2) &= \left\langle \Theta^{-1} z\,,\ H\Theta^{-1} z \right\rangle = ax^2 + 2bxy + cy^2 \\ &= \frac{a+c}{2}|z^2| + \frac{a-c}{2}\mathrm{Re}(z^2) + b\,\mathrm{Im}(z^2). \end{aligned}$$

Hence

$$f(w) = \frac{a+c}{2}|w| + \frac{a-c}{2}\mathrm{Re}(w) + b\,\mathrm{Im}(w)$$

and the result follows. □

Definition 1. *A function $f : \mathbb{C} \to I\!R$ is said to be quadratic on $\mathbb{C}$ exactly when f composed with Θ (defined in (4)) is quadratic on $I\!R^2$.*

We now sharpen the inequality (22) using the notation established at the end of Section 1.

Theorem 4. *If in Theorem 3 it is further assumed that ϕ is either (i) quadratic, or (ii) twice continuously differentiable at λ_0 with $\phi''(\lambda_0; \cdot, \cdot)$ positive definite, then condition (22) can be strengthened to*

$$\eta \geq \frac{1}{n}\left[\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2},\ \sqrt{-b_2})\right]. \tag{29}$$

Proof. If ϕ is quadratic, then the proof follows essentially the same pattern of proof as in the positive definite case. Therefore, we only provide the proof in the case where ϕ is assumed to be twice continuously differentiable at λ_0 with $\phi''(\lambda_0; \cdot, \cdot)$ positive definite.

Let $(v, \eta) \in T_{\mathrm{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$ with v given by (21). By Theorem 3, (v, η) satisfies (22)–(24). By Lemma 1,

$$(\tilde{v}, \eta) \in T_{\mathrm{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right),$$

where $\tilde{v} = \sum_{k=1}^{n} b_k e_{(n-k,\lambda_0)}$. By (24), $b_k = 0,\ k = 3, \ldots, n$ and so there exist sequences $t_r \downarrow 0$, $\eta_r \to \eta$, $b_{1r} \to b_1$, $b_{2r} \to b_2$, and $b_{kr} \to 0,\ k = 3, \ldots, n$ such that

$$\phi(\lambda_0) + \eta_r t_r + o(t_r) \geq \hat{\phi}(p_r), \tag{30}$$

where

$$\begin{aligned} p_r &= e_{(n,\lambda_0)} + \sum_{k=1}^{n}(b_{kr}t_r + o(t_r))e_{(n-r,\lambda_0)} \\ &= e_{(n,\lambda_0)} + (b_{1r}t_r + o(t_r))e_{(n-1,\lambda_0)} + (b_{2r}t_r + o(t_r))e_{(n-2,\lambda_0)} + o(t_r), \end{aligned}$$

with the second equality following since $b_{kr} \to 0,\ k = 3, \dots, n$. Let $\lambda_{1r}, \dots, \lambda_{nr}$ denote the roots of the polynomials p_r for $r = 1, 2, \dots$, respectively. We have

$$\sum_{k=1}^{n} (\lambda_{kr} - \lambda_0) = -b_{1r} t_r + o(t_r), \tag{31}$$

and

$$\begin{aligned}\sum_{k=1}^{n} (\lambda_{kr} - \lambda_0)^2 &= \left[\sum_{k=1}^{n} (\lambda_{kr} - \lambda_0)\right]^2 - 2\sum_{j<k} (\lambda_{jr} - \lambda_0)(\lambda_{kr} - \lambda_0) \\ &= (-b_{1r} t_r + o(t_r))^2 - 2(b_{2r} t_r + o(t_r)) \\ &= -2b_{2r} t_r + o(t_r). \end{aligned} \tag{32}$$

Set $z_{kr} = \lambda_{kr} - \lambda_0$ for $k = 1, \dots, n$ and $r = 1, 2, \dots$. By (31) and (32) we have

$$\sum_{k=1}^{n} z_{kr} = -b_{1r} t_r + o(t_r) \tag{33}$$

and

$$\sum_{k=1}^{n} (z_{kr})^2 = -2b_{2r} t_r + o(t_r), \tag{34}$$

respectively. With this notation, (30) becomes

$$\phi(\lambda_0) + \eta_r t_r + o(t_r) \geq \phi(\lambda_0 + z_{kr}),\ k = 1, \dots, n.$$

Taking second-order Taylor expansions yields

$$\eta_r t_r + o(t_r) \geq \phi'(\lambda_0; z_{kr}) + \frac{1}{2}\phi''(\lambda_0; z_{kr}, z_{kr}) + o(|z_{kr}|^2),\ k = 1, \dots, n, \tag{35}$$

(note that if ϕ is quadratic then the term $o(|z_{kr}|^2)$ equals zero). Since $p_r \to e_{(n,\lambda_0)}$, we have $z_{kr} \to 0,\ k = 1, \dots, n$. Hence, the positive definiteness of $\phi''(\lambda_0; \cdot, \cdot)$ implies that for every $\epsilon > 0$ there is an r_0 such that

$$\frac{1}{2}\phi''(\lambda_0; z_{kr}, z_{kr}) + o(|z_{kr}|^2) \geq \frac{1-\epsilon}{2}\phi''(\lambda_0; z_{kr}, z_{kr}),$$

for $k = 1, \dots, n$ and all $r \geq r_0$ (if ϕ is quadratic then we can take $\epsilon = 0$ without the assumption of positive definiteness). Therefore, for $r \geq r_0$, the inequalities (35) imply the inequalities

$$\eta_r t_r + o(t_r) \geq \phi'(\lambda_0; z_{kr}) + \frac{1-\epsilon}{2}\phi''(\lambda_0; z_{kr}, z_{kr}),\ k = 1, \dots, n.$$

Now sum over k and again use the positive definiteness of $\phi''(\lambda_0;\cdot,\cdot)$ with Lemma 5 and definition (5) to obtain

$$\begin{aligned}\eta_r t_r + o(t_r) &\geq \frac{1}{n}\left[\phi'(\lambda_0;\sum_{k=1}^n z_{kr}) + \frac{1-\epsilon}{2}\sum_{k=1}^n \phi''\left(\lambda_0;\sqrt{z_{kr}^2},\sqrt{z_{kr}^2}\right)\right]\\ &\geq \frac{1}{n}\left[\phi'(\lambda_0;\sum_{k=1}^n z_{kr}) + \frac{1-\epsilon}{2}\phi''\left(\lambda_0;\sqrt{\sum_{k=1}^n z_{kr}^2},\sqrt{\sum_{k=1}^n z_{kr}^2}\right)\right],\end{aligned}$$

for all $r \geq r_0$. Plugging in (33) and (34) gives the relation

$$\eta_r t_r \geq \frac{t_r}{n}\left[\phi'(\lambda_0;-b_{1r}) + (1-\epsilon)\phi''(\lambda_0;\sqrt{-b_{2r}},\sqrt{-b_{2r}})\right] + o(t_r).$$

Dividing through by t_r and taking the limit yields the inequality

$$\eta \geq \frac{1}{n}\left[\phi'(\lambda_0;-b_1) + (1-\epsilon)\phi''(\lambda_0;\sqrt{-b_2},\sqrt{-b_2})\right].$$

Observe that if ϕ is quadratic, we can obtain this inequality with $\epsilon = 0$ without the positive definiteness assumption. Since $\epsilon > 0$ was arbitrary, we obtain (29). □

If ϕ is not quadratic and $\phi''(\lambda_0;\cdot,\cdot)$ is only positive semidefinite, we can still sharpen (22) but not as finely as in (29). The proof in the indefinite case is completely different. Unlike the proof of Theorem 4, it relies only on the Gauss-Lucas Theorem.

Theorem 5. *If in Theorem 3 it is further assumed that ϕ is twice continuously differentiable, then condition (22) can be strengthened to*

$$\eta \geq \frac{1}{n}\left[\phi'(\lambda_0;-b_1) + \frac{1}{(n-1)}\phi''(\lambda_0;\sqrt{-b_2},\sqrt{-b_2})\right], \tag{36}$$

when $n > 1$.

Proof. Consider the polynomial $p = \sum_{k=0}^n a_k e_{(n-k,\lambda_0)}$ and set

$$r = \sqrt{\left(\frac{a_1}{na_0}\right)^2 - \frac{2a_2}{n(n-1)a_0}}$$

The Gauss-Lucas theorem tells us that if $\mu \geq \hat{\phi}(p)$, then

$$\begin{aligned}\mu &\geq \max\left\{\phi(\lambda)\;\middle|\; p^{(n-2)}(\lambda) = 0\right\}\\ &= \max\left\{\phi(\lambda)\;\middle|\;\lambda = \lambda_0 + \frac{-a_1}{na_0} \pm r\right\}\\ &\geq \phi\left(\lambda_0 + \frac{-a_1}{na_0} \pm r\right),\end{aligned}$$

where

$$\begin{aligned}
&\phi\left(\lambda_0 + \frac{-a_1}{na_0} \pm r\right) \\
&\quad = \phi(\lambda_0) + \left\langle \phi'(\lambda_0)\,,\ \frac{-a_1}{na_0}\right\rangle + \frac{1}{2}\left\langle \phi''(\lambda_0)\left(\frac{-a_1}{na_0}\right)\,,\ \left(\frac{-a_1}{na_0}\right)\right\rangle \\
&\qquad \pm \left\langle \phi'(\lambda_0)\,,\ r\right\rangle \pm \left\langle \phi''(\lambda_0)\left(\frac{-a_1}{na_0}\right)\,,\ r\right\rangle \\
&\qquad + \frac{1}{2}\left\langle \phi''(\lambda_0) r\,,\ r\right\rangle + o\left(\left|\frac{-a_1}{na_0} \pm r\right|^2\right).
\end{aligned}$$

By adding the resulting pair of inequalities, one associated with each of the two roots, and then dividing by 2, we get the inequality

$$\begin{aligned}
\mu \geq \phi(\lambda_0) &+ \left\langle \phi'(\lambda_0)\,,\ \frac{-a_1}{na_0}\right\rangle + \frac{1}{2}\left\langle \phi''(\lambda_0)\left(\frac{-a_1}{na_0}\right)\,,\ \left(\frac{-a_1}{na_0}\right)\right\rangle \\
&+ \frac{1}{2}\left\langle \phi''(\lambda_0) r\,,\ r\right\rangle + o\left(\left|\frac{-a_1}{na_0} \pm r\right|^2\right).
\end{aligned} \tag{37}$$

Next, suppose that $(v, \eta) \in T_{\mathrm{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$ with v given by (21). By Lemma 1, $(\tilde{v}, \eta) \in T_{\mathrm{epi}\,(\hat{\phi}_1)}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$, where

$$\tilde{v} = \sum_{k=1}^{n} b_k e_{(n-k,\lambda_0)}.$$

Then, as in the proof of Theorem 3, there exist sequences $t_j \downarrow 0$ and $\{(p_j, \mu_j)\} \in \mathrm{epi}\,(\hat{\phi}_1)$ such that

$$t_j^{-1}((p_j, \mu_j) - (e_{(n,\lambda_0)}, \phi(\lambda_0))) \to (\tilde{v}, \eta).$$

That is, there exists $\{(a_0^j, a_1^j, \ldots, a_n^j)\} \in \mathbb{C}^{n+1}$ such that

$$p_j = \sum_{k=0}^{n} a_k^j e_{(n-k,\lambda_0)},$$

$t_j^{-1}(\mu_j - \phi(\lambda_0)) \to \eta$, $a_0^j = 1$ for all $j = 1, 2, \ldots$, and $t_j^{-1} a_k^j \to b_k$ for $k = 1, \ldots, n$. By replacing μ by μ_j and a_k by a_k^j, $k = 0, \ldots, n$ in (37), dividing through by t_j, and slightly re-arranging, we obtain

$$\begin{aligned}
\frac{\mu_j - \phi(\lambda_0)}{t_j} \geq &\left\langle \phi'(\lambda_0)\,,\ \frac{-a_1^j}{nt_j}\right\rangle + \frac{1}{2}\left\langle \phi''(\lambda_0)\left(\frac{-a_1^j}{nt_j}\right)\,,\ \left(\frac{-a_1^j}{n}\right)\right\rangle \\
&+ \frac{1}{2}\left\langle \phi''(\lambda_0) r^j\,,\ r^j\right\rangle + t_j^{-1} o\left(\left|\frac{-a_1^j}{n} \pm r^j\right|^2\right),
\end{aligned}$$

where

$$r^j = \sqrt{\left(\frac{a_1^j}{nt_j^{\frac{1}{2}}}\right)^2 - \frac{2a_2^j}{n(n-1)t_j}}.$$

Taking the limit in this inequality as $j \to \infty$ yields (36). □

The representation (7) along with Theorems 3, 4, and 5 yield the following representations and bounds for the subderivative of the function $\hat{\phi}$.

Theorem 6. *Let $\hat{\phi}$ be as defined in (1), $\lambda_0 \in \operatorname{dom}(\partial\phi)$ with $\partial\phi(\lambda_0) \neq \{0\}$, and*

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}$$

be given. Then $d\hat{\phi}(e_{(n,\lambda_0)})(v) = +\infty$ if (23) and (24) are not satisfied; otherwise,

$$d\hat{\phi}(e_{(n,\lambda_0)})(v) \geq \frac{1}{n}\phi'(\lambda_0; -b_1), \tag{38}$$

with equality holding if $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$.

If it is further assumed that the function ϕ is twice continuously differentiable at λ_0, then whenever $n > 1$ the inequality (38) can be refined to

$$\frac{1}{n}\left[\phi'(\lambda_0; -b_1) + \frac{1}{n-1}\phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})\right] \tag{39}$$

$$\leq d\hat{\phi}(e_{(n,\lambda_0)})(v) \leq \frac{1}{n}\left[\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})\right], \tag{40}$$

whenever (23) and (24) are both satisfied. "Moreover, quality holds in the second inequality in (40) if any one of the following three conditions hold"

$$\phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) = 0, \tag{41}$$

$$\phi''(\lambda_0; \cdot, \cdot) \text{ is positive definite, or} \tag{42}$$

$$\phi \text{ is quadratic.} \tag{43}$$

Proof. By Theorem 3 we know that the subderivative is $+\infty$ if (23) and (24) are not satisfied. The lower bounds (38) and (39) are immediate consequences of Theorems 3 and 5, respectively.

The representation (7) and Lemma 1 imply that with no loss in generality we may assume for the remainder of the proof that $b_0 = 0$ in v.

Suppose that $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$ and (23) and (24) hold. We show that equality must hold in (38). As noted in Lemma 4, $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$ implies that $b_k = 0,\ k = 2, 3, \ldots, n$. To see that equality is attained consider the family of polynomials

$$\begin{aligned} p_\xi(\lambda) &= (\lambda - \lambda_0 + \xi b_1/n)^n \\ &= (\lambda - \lambda_0)^n + \xi b_1(\lambda - \lambda_0)^{n-1} + o(\xi). \end{aligned}$$

For any sequence of real positive scalars $\{\xi_\nu\}$ decreasing to zero definition (8) shows that

$$\begin{aligned}\phi'(\lambda_0; -b_1/n) &= \lim_{\nu\to\infty} \frac{\phi(\lambda_0 - \xi_\nu b_1/n) - \phi(\lambda_0)}{\xi_\nu} \\ &= \lim_{\nu\to\infty} \frac{\hat{\phi}(p_{\xi_\nu}) - \hat{\phi}(e_{(n,\lambda_0)})}{\xi_\nu} \\ &\geq d\hat{\phi}(e_{(n,\lambda_0)})(v),\end{aligned}$$

hence equality holds in (38).

If either ϕ is quadratic, or $\phi''(\lambda_0; \cdot, \cdot)$ is positive definite, then Theorem 4 tells us that the expression on the right hand side of (40) is also a lower bound. Thus, to show equality in these two cases we need only establish the upper bound (40). In addition, once this upper bound is established then we also obtain equality when (41) holds since in this case the upper bound (40) reduces to the lower bound (38). Thus, it remains only to prove the upper bound (40). We assume throughout that the polynomial v satisfies both (23) and (24).

We use (8) to establish the upper bound (40). The bound is obtained by considering the tangents to smooth curves having as limit $e_{(n,\lambda_0)}$. The proof proceeds by considering the even and odd cases for n separately. But in both cases we make use of the following family of polynomials:

$$\begin{aligned}q_{(\xi,\nu)}(\lambda) &= \left(\lambda - \left(\lambda_0 - \frac{\xi}{n}(b_1 - \frac{1}{2m}\nu) + \sqrt{-b_2\xi/m}\right)\right)^m \\ &\quad \cdot \left(\lambda - \left(\lambda_0 - \frac{\xi}{n}(b_1 - \frac{1}{2m}\nu) - \sqrt{-b_2\xi/m}\right)\right)^m \\ &= \left((\lambda-\lambda_0)^2 + \frac{2\xi}{n}(b_1 - \frac{\nu}{2m})(\lambda-\lambda_0) + b_2\xi/m + o(\xi)\right)^m \\ &= (\lambda-\lambda_0)^{2m} + \frac{2m\xi}{n}(b_1 - \frac{\nu}{2m})(\lambda-\lambda_0)^{2m-1} \\ &\qquad + b_2\xi(\lambda-\lambda_0)^{2m-2} + o(\xi).\end{aligned}$$

First assume that n is even: $n = 2m$ for some positive integer m. Consider the family of polynomials

$$q_{(\xi,0)}(\lambda) = (\lambda-\lambda_0)^n + \xi v(\lambda) + o(\xi).$$

For all ξ, this polynomial has only two roots:

$$\lambda_\xi = \lambda_0 - \frac{b_1}{n}\xi \pm \sqrt{\frac{-b_2}{m}\xi}.$$

For ξ real and positive, the second-order Taylor expansion of ϕ at these roots gives

$$\begin{aligned}\hat{\phi}(q_{(\xi,0)}) &= \max\left\{\phi\left(\lambda_0 - \frac{b_1}{n}\xi + \sqrt{\frac{-b_2}{m}\xi}\right), \phi\left(\lambda_0 - \frac{b_1}{n}\xi - \sqrt{\frac{-b_2}{m}\xi}\right)\right\} \\ &= \phi(\lambda_0) + \xi\left[\phi'(\lambda_0; -b_1/n) + \frac{1}{2}\phi''(\lambda_0; \sqrt{-b_2/m}, \sqrt{-b_2/m})\right] + o(\xi)\end{aligned}$$

since by (23), $\left\langle \phi'(\lambda_0), \sqrt{-b_2} \right\rangle = 0$. Therefore,

$$d\hat{\phi}(e_{(n,\lambda_0)})(v) \le \lim_{\xi \searrow 0} \frac{\hat{\phi}(q_{(\xi,0)}) - \phi(\lambda_0)}{\xi}$$
$$= \frac{1}{n}\left[\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})\right],$$

establishing the even case.

Now consider the odd case with $n = 2m + 1$. This time set

$$\nu = -\frac{\phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})}{\overline{\phi'(\lambda_0)}}$$

and consider the family of polynomials

$$p_\xi(\lambda) = \left(\lambda - \left(\lambda_0 - \frac{\xi}{n}(b_1 + \nu)\right)\right) q_{(\xi,\nu)}(\lambda)$$
$$= (\lambda - \lambda_0)^n + \xi v(\lambda) + o(\xi).$$

For all values of ξ the roots of this polynomial are

$$\lambda_0 - \frac{\xi}{n}(b_1 + \nu) \text{ and } \lambda_0 - \frac{\xi}{n}(b_1 - \frac{1}{2m}\nu) \pm \sqrt{-b_2\xi/m}.$$

Taking the second-order Taylor expansion of ϕ at the root $\lambda_0 - \frac{\xi}{n}(b_1 + \nu)$ for ξ real and positive shows that $\phi(\lambda_0 - \frac{\xi}{n}(b_1 + \nu))$ equals

$$\phi(\lambda_0) + (\xi/n)\left[\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})\right] + o(\xi). \tag{44}$$

Similarly, taking the second-order Taylor expansion of ϕ at either of the two roots $\lambda_0 - \frac{\xi}{n}(b_1 - \frac{1}{2m}\nu) \pm \sqrt{-b_2\xi/m}$ for ξ real and positive and using the fact that $\phi'(\lambda_0; \sqrt{-b_2}) = 0$ shows that

$$\phi\left(\lambda_0 - \frac{\xi}{n}(b_1 - \frac{1}{2m}\nu) \pm \sqrt{-b_2\xi/m}\right)$$

also equals (44). Therefore, for ξ real and positive, we have

$$\hat{\phi}(p_\xi) = \phi(\lambda_0) + \frac{1}{n}(\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}))\xi + o(\xi).$$

The proof is completed as in the even degree case. □

Theorem 3 and its refinements give necessary conditions for inclusion in the tangent cone $T_{\text{epi}\,(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$. We now use the conditions given in Theorem 6 to characterize the tangent cone when ϕ is twice differentiable at λ_0.

Theorem 7. *Let* $\lambda_0 \in \operatorname{dom}(\phi)$ *be such that* $\partial\phi(\lambda_0) \neq \{0\}$*, and set*

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}.$$

If either $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$ *or* ϕ *is twice continuously differentiable at* λ_0 *and any one of the three conditions (41)–(43) hold, then* $(v, \eta) \in T_{\operatorname{epi}(\hat{\phi})}\left(e_{(n,\lambda_0)}, \phi(\lambda_0)\right)$ *if and only if*

$$\begin{aligned} &\eta \geq \frac{1}{n}\left[\phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})\right], \\ &0 = \left\langle g\,,\, \sqrt{-b_2}\right\rangle \ \forall\, g \in \partial\phi(\lambda_0), \quad \textit{and} \\ &0 = b_k,\ k = 3, \ldots, n, \end{aligned} \tag{45}$$

where we interpret the term $\phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})$ *as zero when* ϕ *is not twice continuously differentiable at* λ_0*.*

Proof. Apply Theorem 6 in conjunction with the representation (7). □

3. Regular normals and subgradients

Next consider the variational objects dual to the tangent cone and the subderivative. These are the cone of *regular normals* to the epigraph at a point and the set of *regular subgradients*. The cone of regular normals is the polar of the tangent cone [14, Proposition 6.5]:

$$\begin{aligned} \widehat{N}_{\operatorname{epi}(f)}(x) &= T_{\operatorname{epi}(f)}(x)^{\circ} \\ &= \left\{(z, \tau) \;\middle|\; \langle (z, \tau)\,,\, (w, \mu)\rangle \leq 0,\ \forall\, (w, \mu) \in T_{\operatorname{epi}(f)}(x)\right\}. \end{aligned}$$

A vector v is a regular subgradient [14, Definition 8.3] for f at $x \in \operatorname{dom}(f)$ if

$$f(y) \geq f(x) + \langle v\,,\, y - x\rangle + o(|y - x|). \tag{46}$$

We call the collection of all regular subgradients for f at x the *regular subdifferential* of f at x and denote this set by $\hat{\partial} f(x)$. The regular subdifferential at a point is always a closed convex set. At points x where $\hat{\partial} f(x) \neq \emptyset$ the regular normals and the regular subgradients are related by the formula [14, Theorem 8.9]

$$\widehat{N}_{\operatorname{epi}(f)}(x) = \left\{t(v, -1) \;\middle|\; v \in \hat{\partial} f(x),\ t > 0\right\} \cup \left\{(v, 0) \;\middle|\; v \in \hat{\partial} f(x)^{\infty}\right\}, \tag{47}$$

where $\hat{\partial} f(x)^{\infty}$ denotes the *recession cone* of the set $\hat{\partial} f(x)$. The recession cone of any convex set C is given by

$$C^{\infty} = \{z \mid x + \tau z \in C\ \forall\, x \in C \text{ and } \tau \in I\!R\}.$$

The regular subdifferential is related to the subderivative by the formula [14, Exercise 8.4]

$$\hat{\partial} f(x) = \{v \mid \langle v\,,\, w\rangle \leq df(x)(w)\ \forall\, w\}. \tag{48}$$

Recall that the *support function* for any set D in a Euclidean space E is given by

$$\sigma_D(w) = \sup_{v \in D} \langle v\,,\ w\rangle\,.$$

The support function of a set is a sublinear function and coincides with the support function for the closed convex hull of the set. The domain of the support function for a convex set C is called the *barrier cone* for C, denoted $\mathrm{bar}\,(C)$ [13, Section 13]. The polar of the barrier cone is precisely the recession cone [13, Corollary 14.2.1]:

$$\mathrm{bar}\,(C)^\circ = C^\infty.$$

If the set C is itself a convex cone, then $C^\infty = \mathrm{cl}\,(C)$ and $\mathrm{bar}\,(C) = (C^\infty)^\circ$. Support functions are important in the context of the theory of subdifferentials since the representation (48) implies the inequality

$$\sigma_{\hat{\partial} f(x)}(w) \le df(x)(w) \quad \forall\, w \in E.$$

We use the relation (48) to estimate the regular subdifferential of $\hat{\phi}$ at $e_{(n,\lambda_0)}$ and then use this estimate to approximate the cone of regular normals. Our estimates for the regular subdifferential depends on the following parameterized family of multifunctions $\Delta_\delta\ :\mathrm{dom}\,(\partial\phi) \rightrightarrows \mathbb{C}^{n+1}$ with parameter values $0 \le \delta \in I\!R$. For $\delta = 0$, define $\Delta_0(\lambda_0)$ as the direct product

$$\Delta_0(\lambda_0) = \{0\} \times (-\frac{1}{n}\partial\phi(\lambda_0)) \times \mathcal{K}_{\lambda_0} \times \mathbb{C}^{n-2},$$

where $\mathcal{K}_{\lambda_0}$ is defined in (28). For $\delta > 0$, the multifunction Δ_δ is only defined when ϕ is twice continuously differentiable in which case

$$\Delta_\delta(\lambda_0) = \{0\} \times \{-\phi'(\lambda_0)/n\} \times \Delta_{(\delta,2)}(\lambda_0) \times \mathbb{C}^{n-2},$$

where

$$\Delta_{(\delta,2)}(\lambda_0) = \left\{\theta_2 \;\middle|\; \left\langle \theta_2, \phi'(\lambda_0)^2 \right\rangle \le \delta \left\langle (\mathrm{i}\,\phi'(\lambda_0))\,,\ \phi''(\lambda_0)(\mathrm{i}\,\phi'(\lambda_0)) \right\rangle \right\}.$$

Given $\lambda_0 \in \mathrm{dom}\,(\partial\phi)$ and $\delta \ge 0$, the set $\Delta_\delta(\lambda_0)$, when defined, is a non-empty closed convex set. Since we use (47) to estimate the regular normals from the regular subdifferential, we require expressions for the recession cone of the sets $\Delta_\delta(\lambda_0)$. Since

$$\Delta_{(\delta,2)}(\lambda_0)^\infty = \left\{\theta_2 \;\middle|\; \left\langle \theta_2, \phi'(\lambda_0)^2 \right\rangle \le 0 \right\} = \mathcal{K}_{\lambda_0}$$

and the recession cone of a product of sets is the product of their recession cones, we have

$$\Delta_\delta(\lambda_0)^\infty = \{0\} \times (-\partial\phi(\lambda_0)^\infty) \times \mathcal{K}_{\lambda_0} \times \mathbb{C}^{n-2} \tag{49}$$

for all $\delta \ge 0$.

We now characterize the support function for the sets $\Delta_\delta(\lambda_0)$.

Lemma 6. *Let $\lambda_0 \in \operatorname{dom}(\phi)$ be such that $\partial\phi(\lambda_0) \neq \{0\}$, and let $\delta \geq 0$. Then for every vector*

$$b = (b_0, b_1, \ldots, b_n)^T \in \mathbb{C}^{n+1}$$

we have $\sigma_{\Delta_\delta(\lambda_0)}(b) < +\infty$ if and only if the components of b satisfy (23) and (24), in which case

$$\sigma_{\Delta_\delta(\lambda_0)}(b) = \frac{1}{n}\phi'(\lambda_0; -b_1) + \delta\phi''\left(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}\right). \tag{50}$$

Here the term $\delta\phi''\left(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}\right)$ is to be interpreted as zero when $\delta = 0$ and ϕ is not twice continuously differentiable at λ_0. Finally, if ϕ is twice continuously differentiable at λ_0 with

$$\left\langle (\mathrm{i}\,\phi'(\lambda_0))\,,\ \phi''(\lambda_0)(\mathrm{i}\,\phi'(\lambda_0))\right\rangle = 0,$$

then $\delta\phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) = 0$ for every $b \in \mathbb{C}^{n+1}$ satisfying (23) and (24).

Proof. Since $\Delta_\delta(\lambda_0)$ is a direct product of sets,

$$\begin{aligned}\sigma_{\Delta_\delta(\lambda_0)}(b) &= \sup_{\theta\in\Delta_\delta(\lambda_0)} \langle\theta\,,\ b\rangle \\ &= \frac{1}{n}\phi'(\lambda_0; -b_1) + \sup_{\theta_2\in\Delta_{(\delta,2)}(\lambda_0)} \langle\theta_2\,,\ b_2\rangle + \sum_{j=3}^{n} \sup_{\theta_j\in\mathbb{C}} \left\langle\theta_j\,,\ b_j\right\rangle. \end{aligned} \tag{51}$$

Clearly, $\sigma_{\Delta_\delta(\lambda_0)}(b)$ is finite if and only if both the second and third terms in (51) are finite. We will show that second term is finite if and only if b satisfies (23), and the third term is finite if and only if b satisfies (24).

For the third term in (51) we have $\sum_{j=3}^{n} \sup_{\theta_j\in\mathbb{C}} \left\langle\theta_j\,,\ b_j\right\rangle < \infty$ if and only if $\sup_{\theta_j\in\mathbb{C}} \left\langle\theta_j\,,\ b_j\right\rangle < \infty,\ j = 3, \ldots, n$, or equivalently, (24) holds.

Consider now the second term in (51). It is the support function for the convex set $\Delta_{(\delta,2)}(\lambda_0)$. This term is finite if and only if b_2 is an element of the barrier cone of $\Delta_{(\delta,2)}(\lambda_0)$ [13, page 112]. The barrier cone is contained in the polar of the recession cone of $\Delta_{(\delta,2)}(\lambda_0)$. By (49), this recession cone is the set $\mathcal{K}_{\lambda_0}$, i.e.

$$\operatorname{bar}(\Delta_{(\delta,2)}(\lambda_0)) \subset \mathcal{K}^\circ_{\lambda_0}.$$

In the case where $\delta = 0$, $\Delta_{(0,2)}(\lambda_0)^\infty = \mathcal{K}_{\lambda_0}$, and so $\operatorname{bar}(\Delta_{(0,2)}(\lambda_0)) = \mathcal{K}^\circ_{\lambda_0}$. Therefore, by Lemma 4, $\sup_{\theta_2\in\Delta_{(0,2)}(\lambda_0)} \langle\theta_2\,,\ b_2\rangle < \infty$ if and only if b_2 satisfies (23). For $\delta > 0$, let $b_2 \in \mathcal{K}^\circ_{\lambda_0} = -\operatorname{cone}(\phi'(\lambda_0)^2)$ so that there is a $\tau \in \mathbb{R}_+$ such that $b_2 = \tau\phi'(\lambda_0)^2$. Then for any $\theta_2 \in \Delta_{(\delta,2)}(\lambda_0)$

$$\langle\theta_2\,,\ b_2\rangle = \tau\left\langle\theta_2\,,\ \phi'(\lambda_0)\right\rangle \leq \delta\left\langle(\mathrm{i}\,\phi'(\lambda_0))\,,\ \phi''(\lambda_0)(\mathrm{i}\,\phi'(\lambda_0))\right\rangle < \infty.$$

Hence, $\mathcal{K}^\circ_{\lambda_0} \subset \operatorname{bar}(\Delta_{(\delta,2)}(\lambda_0))$, that is,

$$\operatorname{bar}(\Delta_{(\delta,2)}(\lambda_0)) = \mathcal{K}^\circ_{\lambda_0} \quad \forall\,\delta \geq 0\,,$$

and, again by Lemma 4, $\sup_{\theta_2\in\Delta_{(0,2)}(\lambda_0)} \langle\theta_2\,,\ b_2\rangle < \infty$ if and only if b_2 satisfies (23).

Thus we have shown that $\sigma_{\Delta_\delta(\lambda_0)}(b)$ is finite if and only if the second and third terms in (51) are finite which occurs if and only if b satisfies (23) and (24).

Next suppose that $\sigma_{\Delta_\delta(\lambda_0)}(b)$ is finite, or equivalently, assume that b satisfies (23) (i.e., $b_2 \in \mathcal{K}_{\lambda_0}^\circ$) and (24). We show that (50) holds. If $\delta = 0$, then

$$\begin{aligned}\sup_{\theta_2\in\Delta_{(0,2)}(\lambda_0)} \langle \theta_2\,,\, b_2\rangle &= \sup_{\theta_2\in\mathcal{K}_{\lambda_0}} \langle \theta_2\,,\, b_2\rangle \\ &= 0 \\ &= \delta\left\langle \sqrt{-b_2}\,,\, \phi''(\lambda_0)\sqrt{-b_2}\right\rangle,\end{aligned}$$

where the second equality follows since $b_2 \in \mathcal{K}_{\lambda_0}^\circ$. Now suppose that $\delta > 0$. Then, by Lemma 4(ii), there is a $\tau \in I\!R$ such that $\sqrt{b_2} = \tau\phi'(\lambda_0)$. Hence

$$\begin{aligned}\sup_{\theta_2\in\Delta_{(\delta,2)}(\lambda_0)} \langle \theta_2\,,\, b_2\rangle &= \tau^2 \sup_{\theta_2\in\Delta_{(\delta,2)}} \left\langle \theta_2\,,\, \phi'(\lambda_0)^2\right\rangle \\ &= \tau^2\delta\left\langle (\mathrm{i}\,\phi'(\lambda_0))\,,\, \phi''(\lambda_0)(\mathrm{i}\,\phi'(\lambda_0))\right\rangle \\ &= \delta\left\langle (\mathrm{i}\,\tau\phi'(\lambda_0))\,,\, \phi''(\lambda_0)(\mathrm{i}\,\tau\phi'(\lambda_0))\right\rangle \\ &= \delta\left\langle (\mathrm{i}\sqrt{b_2})\,,\, \phi''(\lambda_0)(\mathrm{i}\sqrt{b_2})\right\rangle \\ &= \delta\left\langle \sqrt{-b_2}\,,\, \phi''(\lambda_0)\sqrt{-b_2}\right\rangle.\end{aligned}$$

Hence, in either case, (50) holds. The final statement of the Theorem also follows from the final derivation above. □

Lemma 6 combined with Theorem 6 and the representation (48) provide a basis for estimates, and in some cases formulas, for the regular subdifferential of $\hat{\phi}$ at $e_{(n,\lambda_0)}$. But first we need to map the sets $\Delta_\delta(\lambda_0)$ into the space of polynomials. Given $\lambda_0 \in \mathbb{C}$ define the linear transformation $\tau_{\lambda_0} : \mathcal{P}^n \to \mathbb{C}^{n+1}$ to be the mapping that takes a polynomial to its Taylor series coefficients when expanded at the base point λ_0, specifically,

$$\tau_{\lambda_0}(p) = \left(p^{(n)}(\lambda_0)/n, \ldots, p'(\lambda_0), p(\lambda_0)\right)^T. \tag{52}$$

Equivalently, if p has the representation $p = \sum_{k=0}^n a_k e_{(n-k,\lambda_0)}$, then $\tau_{\lambda_0}(p) = (a_0, a_1, \ldots, a_n)^T$. The family of linear transformations τ_λ is continuous in λ, and for each λ_0 the transformation τ_{λ_0} is invertible. Indeed, one has

$$\tau_{\lambda_0}^{-1} = \tau_{\lambda_0}^*$$

when the adjoint $\tau_{\lambda_0}^*$ is defined using the inner product $\langle\cdot\,,\,\cdot\rangle_{\lambda_0}$.

Theorem 8. *Let $\hat{\phi}$ be as defined in (1) and $\lambda_0 \in \mathrm{dom}\,(\phi)$ be such that $\partial\phi(\lambda_0) \neq \{0\}$. Then*

$$d\hat{\phi}(e_{(n,\lambda_0)})(v) \geq \sigma_{\Delta_0(\lambda_0)}(\tau_{\lambda_0}(v)) \tag{53}$$

for all $v \in \mathcal{P}^n$ and

$$\tau_{\lambda_0}^*\Delta_0(\lambda_0) \subset \hat{\partial}\hat{\phi}(e_{(n,\lambda_0)}), \tag{54}$$

where $\tau_{\lambda_0}^*$ is the adjoint of τ_{λ_0} with respect to the inner product $\langle \cdot\,,\ \cdot\rangle_{\lambda_0}$. Equality holds in both (53) and (54) if $\mathrm{rspan}\,(\partial\phi(\lambda_0)) = \mathbb{C}$.

If it is further assumed that the function ϕ is twice continuously differentiable at λ_0 with $\phi'(\lambda_0) \neq 0$, then

$$\sigma_{\Delta_{\delta_1}(\lambda_0)}(\tau_{\lambda_0}(v)) \leq d\hat{\phi}(e_{(n,\lambda_0)})(v) \leq \sigma_{\Delta_{\delta_2}(\lambda_0)}(\tau_{\lambda_0}(v)), \tag{55}$$

for all $v \in \mathcal{P}^n$, and

$$\tau_{\lambda_0}^*\Delta_{\delta_1}(\lambda_0) \subset \hat{\partial}\hat{\phi}(e_{(n,\lambda_0)}) \subset \tau_{\lambda_0}^*\Delta_{\delta_2}(\lambda_0) \tag{56}$$

where $\delta_1 = 1/(n(n-1))$ and $\delta_2 = 1/n$. Furthermore, if ϕ is quadratic, or $\phi''(\lambda_0;\cdot,\cdot)$ is positive definite, or $\big\langle (\,\mathrm{i}\,\phi'(\lambda_0))\,,\ \phi''(\lambda_0)(\,\mathrm{i}\,\phi'(\lambda_0))\big\rangle = 0$, then

$$d\hat{\phi}(e_{(n,\lambda_0)})(v) = \sigma_{\Delta_{\delta_2}(\lambda_0)}(\tau_{\lambda_0}(v))$$

for all $v \in \mathcal{P}^n$, and

$$\hat{\partial}\hat{\phi}(e_{(n,\lambda_0)}) = \tau_{\lambda_0}^*\Delta_{\delta_2}(\lambda_0)\ .$$

Proof. Inequality (53) follows from the bound (38) in Theorem 6 coupled with Lemma 6. The left and right inequalities in (55) follow from (39) and (40) in Theorem 6, respectively, again in conjunction with Lemma 6. The subdifferential inclusions in (54) and the left hand side of (56) follow immediately from the definition of the adjoint transformation, the representation (48), the inequalities in (53), and the left hand side of (55), respectively. Here we have also used the identity

$$\sigma_D(A\,\cdot) = \sigma_{A^*D}(\cdot),$$

where A is any linear transformation between Euclidean spaces E_1 and E_2.

The inclusion on the right hand side of (56) follows from the definition of the adjoint transformations and the fact that for any two closed convex sets C_1 and C_2 one has that $C_1 \subset C_2$ if and only if $\sigma_{C_1}(v) \leq \sigma_{C_2}(v)$ for all v.

The final statement of the Theorem follows from the final statement of Theorem 6, the final statement of Lemma 6, and the preceding comment on the relationship between support functions and convex sets. □

We now apply Theorem 8 to (47) obtaining approximations to the cone of regular normals to $\mathrm{epi}\,(\hat{\phi})$ at the point $(e_{(n,\lambda_0)}, \phi(\lambda_0))$. We begin by extending the definitions for the sets $\Delta_\delta(\lambda_0)$ by the formula

$$\Xi_\delta(\lambda_0) = \{\gamma(\theta,-1)\ |\ \theta \in \Delta_\delta(\lambda_0),\ 0 < \gamma \in I\!R\} \cup \Delta_\delta(\lambda_0)^\infty,$$

where the recession cone $\Delta_\delta(\lambda_0)^\infty$ is given by the formula (49). In addition, for each $\lambda \in \mathbb{C}$, define the linear transformation $\hat{\tau}_\lambda\ :\ \mathcal{P}^n \times I\!R \to \mathbb{C}^{n+1} \times I\!R$ by $\hat{\tau}_\lambda(p,\mu) = (\tau_\lambda(p), \mu)$ where the linear transformation τ_λ is defined in (52). The adjoint of $\hat{\tau}_\lambda$, with respect to the inner product

$$\langle (q,\eta)\,,\ (p,\mu)\rangle = \langle q\ ,\ p\rangle_{\lambda_0} + \eta\mu$$

on $\mathcal{P}^n \times I\!R$, is the linear transformation $\hat{\tau}_\lambda^*\ :\ \mathbb{C}^{n+1} \times I\!R \to \mathcal{P}^n \times I\!R$ given by

$$\hat{\tau}_\lambda^*(w,\mu) = (\tau_\lambda^*(w), \mu) = \hat{\tau}_\lambda^{-1}(w,\mu).$$

Theorem 8 and relation (47) give the following corollary to Theorem 8.

Corollary 1. *Let* $\delta \geq 0$ *and* $\hat{\phi}$ *be as defined in (1) with* $\lambda_0 \in \operatorname{dom}(\partial\phi)$ *satisfying* $\partial\phi(\lambda_0) \neq \{0\}$. *Then*

$$\hat{\tau}^*_{\lambda_0}\,\Xi_0(\lambda_0) \subset \widehat{N}_{\operatorname{epi}(\hat{\phi})}(e_{(n,\lambda_0)}, \phi(\lambda_0)), \tag{57}$$

with equality holding if $\operatorname{rspan}(\partial\phi(\lambda_0)) = \mathbb{C}$. *If it is further assumed that the function* ϕ *is twice continuously differentiable at* λ_0 *with* $\phi'(\lambda_0) \neq 0$, *then*

$$\hat{\tau}^*_{\lambda_0}\,\Xi_{\delta_1}(\lambda_0) \subset \widehat{N}_{\operatorname{epi}(\hat{\phi})}(e_{(n,\lambda_0)}, \phi(\lambda_0)) \subset \hat{\tau}^*_{\lambda_0}\,\Xi_{\delta_2}(\lambda_0) \tag{58}$$

where $\delta_1 = 1/(n(n-1))$ *and* $\delta_2 = 1/n$. *Furthermore, if* ϕ *is quadratic, or* $\phi''(\lambda_0;\cdot,\cdot)$ *is positive definite, or* $\big\langle (\mathrm{i}\,\phi'(\lambda_0))\,,\ \phi''(\lambda_0)(\mathrm{i}\,\phi'(\lambda_0))\big\rangle = 0$, *then*

$$\widehat{N}_{\operatorname{epi}(\hat{\phi})}(e_{(n,\lambda_0)}, \phi(\lambda_0)) = \hat{\tau}^*_{\lambda_0}\,\Xi_{\delta_2}(\lambda_0)\ .$$

4. The abscissa mapping

We apply the results of the preceding two sections to the abscissa mapping for polynomials:

$$\mathsf{a}(p) = \sup\{\operatorname{Re}\lambda \mid \lambda \in \mathcal{R}(p)\}\,.$$

Here $\mathsf{a} = \hat{\phi}$, where $\hat{\phi}$ is defined in (1) with the function ϕ given by the linear form

$$\phi(\lambda) = \langle 1\,,\ \lambda\rangle\,.$$

Since $\phi''(\lambda) \equiv 0$, we obtain complete characterizations for the variational objects under study. We state two results. The first concerns the tangent cone and subderivative, and the second the regular normals and subdifferential.

Theorem 9. *Given* $\lambda_0 \in \mathbb{C}$, *one has* $(v, \eta) \in T_{\operatorname{epi}(\mathsf{a})}\big(e_{(n,\lambda_0)}, \operatorname{Re}(\lambda_0)\big)$, *with*

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}$$

if and only if

$$\begin{aligned} &\eta \geq -\frac{1}{n}\operatorname{Re}(b_1),\\ &0 \leq \operatorname{Re} b_2, \quad 0 = \operatorname{Im} b_2, \quad \textit{and}\\ &0 = b_k,\ k = 3, \ldots, n. \end{aligned}$$

Moreover, $d\mathsf{a}(e_{(n,\lambda_0)})(v) = +\infty$ *if (59) and (59) are not satisfied; otherwise,*

$$d\mathsf{a}(e_{(n,\lambda_0)})(v) = -\frac{1}{n}\operatorname{Re}(b_1).$$

Proof. The final statement of the theorem follows immediately from the final statement of Theorem 6. The first part of the result follows from Theorem 7. □

Remark 2. The two conditions $0 \leq \operatorname{Re} b_2$ and $0 = \operatorname{Im} b_2$ in (59) are equivalent to the single condition $0 = \operatorname{Re} \sqrt{-b_2}$ which follows from condition (45) in Theorem 7.

The subdifferential and normal cone characterizations follow directly from Theorem 8 and Corollary 1.

Theorem 10. *Let $\lambda_0 \in \mathbb{C}$ be given. Then*

$$\widehat{N}_{\mathrm{epi}\,(\mathsf{a})}(e_{(n,\lambda_0)}, \operatorname{Re}(\lambda_0)) = \hat{\tau}^*_{\lambda_0} \left\{ t(w,-1) \;\middle|\; \begin{array}{c} 0 \leq t,\ w_0 = 0,\ w_1 = \frac{-1}{n} \\ \textit{and} \operatorname{Re}(w_2) \leq 0 \end{array} \right\},$$

and

$$\partial \mathsf{a}(e_{(n,\lambda_0)}) = \tau^*_{\lambda_0} \{ w \mid w_0 = 0,\ w_1 = -1/n,\ \textit{and} \operatorname{Re}(w_2) \leq 0 \}.$$

The results of Theorems 9 and 10 coincide precisely with those found in [3, 5].

5. The radius mapping

Consider now the radius mapping for polynomials:

$$\mathsf{r}(p) = \sup \{ |\lambda| \mid \lambda \in \mathcal{R}(p) \}.$$

Here $\mathsf{r} = \hat{\phi}$, where $\hat{\phi}$ is defined in (1) with the function ϕ given by the modulus

$$\phi(\lambda) = |\lambda|.$$

The modulus is convex and infinitely differentiable in the real sense except at the origin. The convex subdifferential is given by

$$\partial\, |\cdot|\, (\zeta) = \begin{cases} \mathsf{B}, & \text{if } \zeta = 0; \\ \zeta / |\zeta|, & \text{otherwise,} \end{cases}$$

where $\mathsf{B} = \{ \zeta \mid |\zeta| \leq 1 \}$ is the closed unit disk in $\mathbb{C}$. At nonzero ζ we have

$$|\cdot|''(\zeta; \delta, \delta) = \frac{1}{|\zeta|} \left[|\delta|^2 - \langle \zeta / |\zeta|\,,\, \delta \rangle^2 \right].$$

Since for $\lambda_0 \neq 0$ the Hessian is not positive definite and

$$\frac{1}{|\lambda_0|} = \left\langle \left(\frac{\mathrm{i}\,\lambda_0}{|\lambda_0|} \right),\ \phi''(\lambda_0) \left(\frac{\mathrm{i}\,\lambda_0}{|\lambda_0|} \right) \right\rangle,$$

it would seem that our strongest results for the polynomial $e_{(\lambda_0,n)}$ do not apply when $\lambda_0 \neq 0$. However, this difficulty is easily sidestepped.

Lemma 7. *Let $p \in \mathcal{M}^n$ be any polynomial for which $\mathsf{r}(p) > 0$. Then $(v,\eta) \in T_{\mathrm{epi}\,(\mathsf{r})}(p,\mu)$ if and only if $(v,\mu\eta) \in T_{\mathrm{epi}\,(\mathsf{r}_2)}(p, \frac{1}{2}\mu^2)$, where*

$$\mathsf{r}_2(p) = \sup \left\{ \frac{1}{2} |\lambda|^2 \mid \lambda \in \mathcal{R}(p) \right\}.$$

Proof. Let $(v, \eta) \in T_{\text{epi}\,(\mathsf{r})}(p, \mu)$. Then there exist sequences

$$\{(p_k, \mu_k)\} \subset \text{epi}\,(\mathsf{r}) \text{ and } t_k \searrow 0 \tag{59}$$

such that

$$\frac{p_k - p}{t_k} \to v, \text{ and} \tag{60}$$

$$\frac{\mu_k - \mu}{t_k} \to \eta. \tag{61}$$

Moreover, we may assume with no loss in generality that $\mu_k > 0$ for all k since $\mu \geq \mathsf{r}(p) > 0$.

Now since $(p, \mu) \in \text{epi}\,(\mathsf{r})$ if and only if $(p, \mu^2/2) \in \text{epi}\,(\mathsf{r}_2)$, we have (59) is equivalent to

$$\{(p_k, \mu_k^2/2)\} \subset \text{epi}\,(\mathsf{r}_2) \text{ and } t_k \searrow 0. \tag{62}$$

Also, since $0 < \mu, \mu_k$, (61) is equivalent to

$$\frac{(\mu_k - \mu)(\mu_k + \mu)}{t_k} \to 2\mu\eta$$

or equivalently,

$$\frac{\frac{1}{2}\mu_k^2 - \frac{1}{2}\mu^2}{t_k} \to \mu\eta. \tag{63}$$

Therefore, the statements (59)–(61) are equivalent to the statements (62), (60), and (63), or equivalently, $(v, \mu\eta) \in T_{\text{epi}\,(\mathsf{r}_2)}(p, \mu)$. □

Lemma 7 gives the representation

$$T_{\text{epi}\,(\mathsf{r})}(p, \mu) = \left\{ (v, \eta/\mu) \;\middle|\; (v, \eta) \in T_{\text{epi}\,(\mathsf{r}_2)}(p, \mu^2/2) \right\}, \tag{64}$$

whenever $\mathsf{r}(p) > 0$. Since $\frac{1}{2}\,|\cdot|^2$ is quadratic, with $(\frac{1}{2}\,|\cdot|^2)''(\zeta; \delta, \delta) = |\delta|^2$, Theorem 7 provides a complete characterization of the tangent cone $T_{\text{epi}\,(\mathsf{r})}(e_{(n,\lambda_0)}, |\lambda_0|)$.

Theorem 11. *Let $\lambda_0 \in \mathbb{C}$ and let $(v, \eta) \in \mathcal{P}^n \times I\!R$ be such that*

$$v = \sum_{k=0}^{n} b_k e_{(n-k,\lambda_0)}.$$

(i) If $\lambda_0 = 0$, then $(v, \eta) \in T_{\text{epi}\,(\mathsf{r})}(e_{(n,0)}, \lambda_0)$ if and only if

$$\eta \geq \frac{1}{n}\,|b_1|,$$
$$0 = b_k,\; k = 2, 3, \ldots, n.$$

Moreover, $d\mathsf{r}(e_{(n,\lambda_0)})(v) = +\infty$ if (65) is not satisfied; otherwise,

$$d\mathsf{r}(e_{(n,\lambda_0)})(v) = \frac{1}{n}\,|b_1|.$$

(ii) If $\lambda_0 \neq 0$*, then* $(v, \eta) \in T_{\mathrm{epi}\,(\mathrm{r})}(e_{(n,\lambda_0)}, |\lambda_0|)$ *if and only if*

$$\begin{aligned} \eta &\geq \frac{1}{n\,|\lambda_0|}\left[|b_2| - \operatorname{Re}\bar{\lambda}_0 b_1\right], \\ 0 &= \operatorname{Re}\bar{\lambda}_0\sqrt{-b_2}, \ \textit{and} \\ 0 &= b_k,\ k = 3, \ldots, n. \end{aligned}$$

Moreover, $d\mathrm{r}(e_{(n,\lambda_0)})(v) = +\infty$ *if (65) and (65) are not satisfied; otherwise,*

$$d\mathrm{r}(e_{(n,\lambda_0)})(v) = \frac{1}{n\,|\lambda_0|}\left[|b_2| - \operatorname{Re}\bar{\lambda}_0 b_1\right].$$

Proof. The case $\lambda_0 = 0$ follows from Theorem 6 and the representation (7) since the fact that $\mathsf{B} = \partial\,|\cdot|\,(0)$ has non-empty interior implies that $\operatorname{rspan}(\partial\,|\cdot|\,(0)) = \mathbb{C}$. The case $\lambda_0 \neq 0$ follows from Theorem 7 and the representation (64). $\square$

We have the following dual variational results for the regular subdifferential and normal cone.

Theorem 12. *Let* $\lambda_0 \in \mathbb{C}$ *and let the linear transformation* $\tau_\lambda : \mathcal{P}^n \to \mathbb{C}^{n+1}$ *be as defined in (52). Set*

$$\Delta^{\mathrm{r}}(\lambda_0) = \left\{ \begin{pmatrix} \theta_0 \\ \vdots \\ \theta_n \end{pmatrix} \,\middle|\, \theta_0 = 0,\ \theta_1 \in \frac{1}{n}\mathsf{B} \right\},$$

if $\lambda_0 = 0$*; otherwise, set*

$$\Delta^{\mathrm{r}}(\lambda_0) = \left\{ \begin{pmatrix} \theta_0 \\ \vdots \\ \theta_n \end{pmatrix} \,\middle|\, \theta_0 = 0,\ \theta_1 = \frac{-\lambda_0}{n\,|\lambda_0|},\ \left\langle \theta_2\,,\ \lambda_0{}^2 \right\rangle \leq \frac{|\lambda_0|}{n} \right\}.$$

(i) If $\lambda_0 = 0$*, then*

$$\begin{aligned} d\mathrm{r}(e_{(n,\lambda_0)})(v) &= \sigma_{\Delta^{\mathrm{r}}(\lambda_0)}(\tau_{\lambda_0}(v)), \\ \hat{\partial}\mathrm{r}(e_{(n,\lambda_0)}) &= \tau_{\lambda_0}^{*}\Delta^{\mathrm{r}}(\lambda_0), \end{aligned}$$

and

$$\widehat{N}_{\mathrm{epi}\,(\mathrm{r})}(e_{(n,\lambda_0)}, 0) = \left\{ (\tau_{\lambda_0}^{*}(w), -\mu) \,\middle|\, \begin{array}{c} \mu \geq 0,\ w \in \mathbb{C}^{n+1}, \\ w_0 = 0,\ |w_1| \leq \mu \end{array} \right\}.$$

(ii) If $\lambda_0 \neq 0$*, then*

$$d\mathrm{r}(e_{(n,\lambda_0)})(v) = \sigma_{\Delta^{\mathrm{r}}(\lambda_0)}(\tau_{\lambda_0}(v)), \tag{65}$$

$$\hat{\partial}\mathrm{r}(e_{(n,\lambda_0)}) = \tau_{\lambda_0}^{*}\Delta^{\mathrm{r}}(\lambda_0), \tag{66}$$

and

$$\widehat{N}_{\mathrm{epi}\,(\mathrm{r})}(e_{(n,\lambda_0)}, \lambda_0) = \left\{ (\tau_{\lambda_0}^{*}(w), -\mu) \,\middle|\, \begin{array}{c} \mu \geq 0,\ w_0 = 0,\ w_1 = \frac{-\mu\lambda_0}{|\lambda_0|}, \\ \left\langle w_2\,,\ \lambda_0{}^2 \right\rangle \leq \mu\,|\lambda_0| \end{array} \right\}. \tag{67}$$

Proof. Let us first suppose that $\lambda_0 = 0$. In this case, $\text{rspan}\,(\partial\,|\cdot|\,(0)) = \mathbb{C}$ since the subdifferential $\mathsf{B} = \partial\,|\cdot|\,(0)$ has non-empty interior, hence the results of Theorems 6 and (58) directly apply to give the result.

Next suppose that $\lambda_0 \neq 0$. In this case, Lemma 6 combined with Part (2) of Theorem 11 gives (65). This in turn establishes (66) due to the relation (48). The final relation (67) follows from the equivalence (47). □

6. Concluding remarks

We have shown that the Gauss-Lucas technique presented in [3] extends nicely to the class (1) obtaining first-order necessary condition for inclusion in the tangent cone $T_{e_{(n,\lambda_0)}}(\text{epi}\,(\hat{\phi}))$ (Theorem 3). However, substantial additional work was required to obtain the second-order necessary and sufficient conditions given in Theorem 7. It is gratifying that the second-order result preserves the simplicity and geometric appeal of Theorem 3. Simply stated the result says that first-order growth in $\hat{\phi}$ is controlled not only by $\phi'(\lambda_0)$ but also by the second-order behavior in directions that are both perpendicular to $\phi'(\lambda_0)$ and correspond to a square root splitting of the roots. This is illustrated in the application to the radius mapping in Section 5.

Regrettably, Theorem 7 is still incomplete in the case where $\phi''(\lambda_0)$ is indefinite. We conjecture that the result continues to hold in this case. This conjecture is closely related to the much deeper conjecture that the functions $\hat{\phi}$ are prox-regular [14, Definition 13.27] at points where ϕ is twice differentiable. If true, this result would make a number of results possible including the extension of Theorem 7 to the indefinite case. Indeed, the prox-regularity question is at this time the most important unresolved issue concerning this class of functions.

The extension of the results in the previous sections to general polynomials on $\mathcal{M}^n$ and issues of subdifferential regularity are part of our ongoing work.

Acknowledgements. We wish to thank two anonymous referees for their careful reading of the original manuscript. Their efforts have contributed greatly to the quality of the paper. We would also like to thank Boris Mordukhovich for all of his efforts as Editor of this special issue.

References

1. Burke, J.V., Lewis, A.S., Overton, M.L.: Optimizing matrix stability. Proc. of the American Math. Soc. **129**, 1635–1642 (2000)
2. Burke, J.V., Lewis, A.S., Overton, M.L.: Optimal stability and eigenvalue multiplicity. Foundations of Computational Mathematics **1**, 205–225 (2001)
3. Burke, J.V., Lewis, A.S., Overton, M.L.: Variational analysis of the abscissa mapping for polynomials via the Gauss-Lucas theorem. Journal of Global Optimization **28**, 259–268 (2004)
4. Burke, J.V., Overton, M.L.: Variational analysis of the abscissa mapping for polynomials. SIAM Journal on Control and Optimization **39**, 1651–1676 (2000)
5. Burke, J.V., Overton, M.L.: Variational analysis of non–Lipschitz spectral functions. Math. Programming **90**, 317–351 (2001)
6. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York 1983. Republished by SIAM, 1990
7. Clarke, F.H., Ledyaev, Yu.S., Stern, R.J., Wolenski, P.R.: Nonsmooth Analysis and Control Theory. Springer, 1998

8. Levantovskii, L.V.: On singularities of the boundary of a stability region. Moscow Univ. Math. Bull. **35**, 19–22 (1980)
9. Lucas, F.: Sur une application de la méchanique rationnelle á la théorie des équations. C.R. Acad. Sci. Paris **89**, 224–226 (1879)
10. Lucas, F.: Géométrie des Polynômes. J. École Polytech **46**, 1–33 (1879)
11. Marden, M.: Geometry of Polynomials. American Mathematical Society, 1966
12. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation, Vol. 1: Basic Theory, Vol. 2: Applications. Springer, 2005
13. Rockafellar, R.T.: Convex Analysis. Princeton University Press, 1970
14. Rockafellar, R.T., Wets, R.J-B.: Variational Analysis. Springer, 1998