

# VARIATIONAL ANALYSIS OF FUNCTIONS OF THE ROOTS OF POLYNOMIALS

JAMES V. BURKE<sup>1</sup>, ADRIAN S. LEWIS<sup>2</sup>, AND MICHAEL L. OVERTON<sup>3</sup>

ABSTRACT. The Gauss-Lucas Theorem on the roots of polynomials nicely simplifies calculating 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.

---

*Date:* June 29, 2004.

1. Research supported in part by the National Science Foundation Grant DMS-0203175.
2. Research supported in part by the Natural Sciences and Engineering Research Council of Canada.
3. Research supported in part by the National Science Foundation Grant CCR-0098145.

## 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  $\mathcal{R} : \mathcal{P}^n \rightarrow \mathbb{C}$  given by

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

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

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

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 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 focus our study of the class (1) to their behavior 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  $p_\epsilon(\lambda) = (\lambda - \lambda_0)^n - \epsilon$ . For the unboundedness of  $\hat{\phi}$  on the boundary of  $\mathcal{M}^n$  recall that for any non-constant convex function  $\phi$  there must exist a nonzero direction  $a \in \mathbb{C}$  for which  $\phi(\tau a) \rightarrow +\infty$  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  $q_\epsilon \rightarrow p$  as  $\epsilon \searrow 0$ . Moreover,  $\epsilon^{-1}a \in \mathcal{R}(q_\epsilon)$  for all  $\epsilon > 0$ . Therefore,  $\hat{\phi}(q_\epsilon) \rightarrow +\infty$  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, 2, 3, 4, 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 1.1.** *[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  $n$ :

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

where  $\text{conv } S$  denotes the convex hull of the set  $S$ .

Theorem 1.1 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].

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 1.2.** Let  $r : \mathcal{P}^n \rightarrow \mathbb{R}$  denote the radius mapping on  $\mathcal{P}^n$ :

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

Let  $(v, \eta) \in \mathcal{P}^n \times \mathbb{R}$  be such that

$$(3) \quad v = \sum_{k=0}^n b_k e_{(n-k, \lambda_0)}.$$

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

$$(4) \quad \eta \geq \frac{1}{n} |b_1|,$$

$$(5) \quad 0 = b_k, \quad k = 2, 3, \dots, n.$$

(ii)  $(v, \eta)$  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

$$(6) \quad \eta \geq \frac{1}{n |\lambda_0|} [ |b_2| - \operatorname{Re} \bar{\lambda}_0 b_1 ],$$

$$(7) \quad 0 = \operatorname{Re} \bar{\lambda}_0 \sqrt{-b_2}, \text{ and}$$

$$(8) \quad 0 = b_k, \quad k = 3, \dots, n.$$

The notation and definitions follow those established in [14]. The definitions of the terms *epigraph* and *tangent cone* used in Theorem 1.2 appear in the next section.

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

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

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

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

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

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

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  $\lambda_0$

in the sense that the mapping  $(p, q, \lambda_0) \rightarrow \langle p, q \rangle_{\lambda_0}$  is continuous on  $\mathcal{P}^n \times \mathcal{P}^n \times \mathbb{C}$ . To see this simply note that

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

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  $\phi : \mathbb{C} \rightarrow \overline{\mathbb{R}}$ , we define the mapping  $\tilde{\phi} : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$  by the composition  $\tilde{\phi} = \phi \circ \Theta$ , where  $\Theta : \mathbb{R}^2 \rightarrow \mathbb{C}$  is the linear transformation

$$(9) \quad \Theta(x) = x_1 + i x_2,$$

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

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

We say that  $\phi$  is *differentiable in the real sense* if  $\tilde{\phi}$  is differentiable, in which case the derivative of  $\phi$  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  $\phi$  is twice differentiable in the real sense if  $\tilde{\phi}$  is twice differentiable and again the chain rule gives

$$(10) \quad \phi''(\zeta) \delta = \Theta \nabla^2 \tilde{\phi}(\Theta^* \zeta) \Theta^* \delta,$$

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  $\overline{\mathbb{R}}$  is differentiable or twice differentiable. We also make use of the following notation:

$$\begin{aligned} \phi'(\zeta; \delta) &= \langle \phi'(\zeta), \delta \rangle, \text{ and} \\ \phi''(\zeta; \omega, \delta) &= \langle \omega, \phi''(\zeta) \delta \rangle. \end{aligned}$$

## 2. THE TANGENT CONE AND THE SUBDERIVATIVE

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

$$(11) \quad \operatorname{epi}(f) = \{(x, \mu) \mid f(x) \leq \mu\}.$$

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

$$T_C(x) = \left\{ d \mid \begin{array}{l} \exists \{x^k\} \subset C, \{t_k\} \subset \mathbb{R} \text{ such that} \\ x^k \rightarrow x, t_k \downarrow 0, \text{ and } t_k^{-1}(x^k - x) \rightarrow d \end{array} \right\}.$$

The tangent cone to  $\text{epi}(f)$  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]:

$$(12) \quad \text{epi}(df(x)) = \{(w, \eta) \mid df(x)(w) \leq \eta\} = T_{\text{epi}(f)}(x, f(x))$$

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]:

$$(13) \quad df(x)(w) = \liminf_{\substack{t \searrow 0 \\ w' \rightarrow w}} \frac{f(x + tw') - f(x)}{t}.$$

In the case of  $\hat{\phi}$  defined in (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  $\zeta$ .

**Lemma 2.1.** *Define  $\mathcal{M}_1^n \subset \mathcal{M}^n$  to be the set of monic polynomials of degree  $n$ :*

$$\mathcal{M}_1^n = \{e_{(n,0)} + q \mid q \in \mathcal{P}^{n-1}\}.$$

Define  $\hat{\phi}_1 : \mathcal{P}^n \rightarrow \bar{\mathbb{R}}$  by

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

Let  $\lambda_0 \in \text{dom}(\phi)$ ,  $b_k \in \mathbb{C}$ ,  $k = 0, 1, \dots, n$ ,  $\eta \in \mathbb{R}$ , 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

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

if and only if

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

**Remark** The lemma shows that

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

where  $\mathbb{C}e_{(n, \lambda_0)} = \{\xi e_{(n, \lambda_0)} \mid \xi \in \mathbb{C}\}$ . 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  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0))$ , 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,

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

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

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

$$(14) \quad (e_{(n, \lambda_0)} + \xi_j \tilde{v} + o(\xi_j), \phi(\lambda_0) + \xi_j \eta + o(\xi_j)) \in \text{epi}(\hat{\phi}_1).$$

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

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

Hence (14) 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,  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0))$  which proves the result.  $\square$

Next recall that the epigraph of the convex function  $\phi$  as well as all of the lower level sets

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

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

$$(15) \quad \mu \geq \hat{\phi}(p) \iff \mathcal{R}(p) \subset \text{lev}_\phi(\mu),$$

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

$$(16) \quad \text{conv } \mathcal{R}(p^{(n-1)}) \subset \cdots \subset \text{conv } \mathcal{R}(p') \subset \text{conv } \mathcal{R}(p) \subset \text{lev}_\phi(\mu).$$

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

The convexity of  $\phi$  implies that  $\phi$  is directionally differentiable in all directions at every point in  $\lambda_0 \in \text{dom}(\phi)$  [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  $\omega$  is a subgradient of  $\phi$ , written  $\omega \in \partial\phi(\lambda)$  [13], if and only if

$$(17) \quad \phi(\lambda) \geq \phi(\lambda_0) + \langle \omega, \lambda - \lambda_0 \rangle \quad \forall \lambda \in \mathbb{C}.$$

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  $\text{dom}(\phi)$ . The subdifferential is related to the directional derivative of  $\phi$  by the formula

$$(18) \quad \phi'(\lambda_0; \lambda) = \sup \{ \langle z, \lambda \rangle \mid z \in \partial\phi(\lambda_0) \}.$$

Therefore, at points

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

we have  $\phi'(\lambda_0; \cdot) : \mathbb{C} \rightarrow \mathbb{R} \cup \{+\infty\}$  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  $\hat{\phi}_{\lambda_0} : \mathcal{P}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$  by

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

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

**Lemma 2.2.** *Given  $0 < \tau \in \mathbb{R}$  and  $\lambda_0 \in \text{dom}(\phi)$ , define the linear transformation  $\mathcal{S}_{[\tau, \lambda_0]} : \mathcal{P}^n \rightarrow \mathcal{P}^n$  by*

$$\mathcal{S}_{[\tau, \lambda_0]}(p)(\lambda) = p(\lambda_0 + \tau\lambda).$$

*Then, for every  $p \in \mathcal{P}^n$ ,*

$$(19) \quad \hat{\phi}(p) \geq \phi(\lambda_0) + \tau \hat{\phi}_{\lambda_0}([\mathcal{S}_{[\tau, \lambda_0]}(p)]^{(\ell)}),$$

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

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

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

The remaining cases follow immediately from the the remark preceding the proof, the equivalence (15), and the chain of inclusions (16).  $\square$



We now translate the content of Lemma 2.2 into statements about the coefficients of the underlying polynomials. Consider the polynomial

$$(20) \quad p = \sum_{k=0}^n a_k e_{(n-k, \lambda_0)},$$

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

$$(21) \quad p^{(\ell)} = \ell! \sum_{k=0}^{n-\ell} b(n-k, \ell) a_k e_{(n-(k+\ell), \lambda_0)},$$

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

$$(22) \quad \mathcal{S}_{[\tau, \lambda_0]}(p) = \sum_{k=0}^n a_k \tau^{(n-k)} e_{(n-k, 0)},$$

and

$$(23) \quad [\mathcal{S}_{[\tau, \lambda_0]}(p)]^{(\ell)} = \ell! \tau^n \sum_{k=0}^{n-\ell} b(n-k, \ell) \tau^{-k} a_k e_{(n-(k+\ell), 0)},$$

for  $\ell = 0, 1, 2, \dots, (n-1)$  (the case  $\ell = 0$  is just (22)). With this notation, we have the following simple consequence of Lemma 2.2.

**Lemma 2.3.** *Let  $p \in \mathcal{P}^n$  be as in (20) and let  $t \in \mathbb{R}$  be positive. Then*

$$(24) \quad \hat{\phi}(p) \geq \phi(\lambda_0) + t^{1/s} \hat{\phi}_{\lambda_0} \left( \sum_{k=0}^s b(n-k, n-s) t^{-k/s} a_k e_{(s-k, 0)} \right),$$

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

*Proof.* In (23) take  $\ell = n-s$  and  $\tau = t^{1/s}$  for  $\ell = 0, 1, 2, \dots, (n-1)$ , or equivalently,  $s = 1, \dots, n$ , 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  $s = 1, \dots, n$ . Plugging this expression into (19) yields the result since  $\hat{\phi}_{\lambda_0}(p) = \hat{\phi}_{\lambda_0}(\zeta p)$  for every nonzero complex number  $\zeta$ .  $\square$

The main result of the section now follows.

**Theorem 2.4.** *Let  $\lambda_0 \in \text{dom}(\partial\phi)$  with  $\partial\phi(\lambda_0) \neq \{0\}$ . If  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}((e_{(n, \lambda_0)}, \phi(\lambda_0)))$  with*

$$(25) \quad v = \sum_{k=0}^n b_k e_{(n-k, \lambda_0)},$$

then

$$(26) \quad \eta \geq \phi'(\lambda_0; -b_1/n),$$

$$(27) \quad 0 = \langle g, \sqrt{-b_2} \rangle \quad \forall g \in \partial\phi(\lambda_0), \quad \text{and}$$

$$(28) \quad 0 = b_k, \quad k = 3, \dots, n.$$

Thus, in particular, if  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$ , then  $b_2 = 0$ .

*Proof.* Let  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0))$  with  $v$  given by (25). By Lemma 2.1,  $(\tilde{v}, \eta) \in T_{\text{epi}(\hat{\phi}_1)}(e_{(n, \lambda_0)}, \phi(\lambda_0))$ , where

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

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

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

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

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

$t_j^{-1}(\mu_j - \phi(\lambda_0)) \rightarrow \eta$ ,  $a_0^j = 1$  for all  $j = 1, 2, \dots$ , and  $a_k^j \rightarrow 0$  with  $t_j^{-1}a_k^j \rightarrow b_k$  for  $k = 1, \dots, n$ . By applying Lemma 2.3 to  $p_j$  for each  $j = 1, 2, \dots$  with  $t = t_j$  we obtain for  $s = 1, 2, \dots, n$

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

or equivalently,

$$(29) \quad t_j^{-1/s}(\mu_j - \phi(\lambda_0)) \geq \hat{\phi}_{\lambda_0} \left( \sum_{k=0}^s b(n-k, n-s) t_j^{-k/s} a_k^j e_{(s-k, 0)} \right)$$

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

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

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

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

for  $s = 1, \dots, n$ , since  $a_0^j = 1$  for all  $j = 1, 2, \dots$  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 (29) imply that

$$(30) \quad \eta \geq \hat{\phi}_{\lambda_0} (b(n, n-1) e_{(1,0)} + b_1) = \phi'(\lambda_0; -b_1/n)$$

and

$$(31) \quad \begin{aligned} 0 &\geq \hat{\phi}_{\lambda_0} (b(n, n-s) e_{(s,0)} + b_s) \\ &= \max \left\{ \phi' \left( \lambda_0; \left( \frac{-b_s}{b(n, n-s)} \right)^{1/s} \omega \right) \mid \begin{array}{l} \omega = e^{2\pi k i / s} \\ k = 0, \dots, s-1 \end{array} \right\} \end{aligned}$$

for  $s = 2, \dots, n$ . This proves (26). By assumption there exists  $g \in \partial\phi(\lambda_0)$  with  $g \neq 0$ . Inequality (31) implies that

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

for  $\omega = e^{2\pi k i / s}$  ( $k = 0, 1, \dots, s-1$ ). For  $s = 3, \dots, n$  this can only occur if  $b_s = 0$  which gives (28). For  $s = 2$  we have

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

or equivalently, condition (27) holds.  $\square$

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

**Lemma 2.5.** *If  $H$  is a 2-by-2 real symmetric matrix with nonnegative trace then the function*

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

*is sublinear, i.e. positive homogeneous and subadditive (see (9) for the definition of  $\Theta$ ).*

*Proof.* If

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

and  $z = x + iy$  with  $x$  and  $y$  real, then

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

Hence

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

and the result follows.  $\square$

**Definition 2.6.** A function  $f : \mathbb{C} \rightarrow \mathbb{R}$  is said to be quadratic on  $\mathbb{C}$  exactly when  $f$  composed with  $\Theta$  (defined in (9)) is quadratic on  $\mathbb{R}^2$ .

We now sharpen the inequality (26) using the notation established at the end of Section 1.

**Theorem 2.7.** If in Theorem 2.4 it is further assumed that  $\phi$  is either (i) quadratic, or (ii) twice continuously differentiable at  $\lambda_0$  with  $\phi''(\lambda_0; \cdot, \cdot)$  positive definite, then condition (26) can be strengthened to

$$(32) \quad \eta \geq \frac{1}{n} \left[ \phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) \right].$$

*Proof.* If  $\phi$  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  $\phi$  is assumed to be twice continuously differentiable at  $\lambda_0$  with  $\phi''(\lambda_0; \cdot, \cdot)$  positive definite.

Let  $(v, \eta) \in T_{\operatorname{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0))$  with  $v$  given by (25). By Theorem 2.4,  $(v, \eta)$  satisfies (26)-(28). By Lemma 2.1,

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

where  $\tilde{v} = \sum_{k=1}^n b_k e_{(n-k, \lambda_0)}$ . From (27) and (28), there exist sequences  $t_r \downarrow 0$ ,  $\eta_r \rightarrow \eta$ ,  $b_{1r} \rightarrow b_1$ , and  $b_{2r} \rightarrow b_2$  such that

$$(33) \quad \phi(\lambda_0) + \eta_r t_r + o(t_r) \geq \hat{\phi}(p_r),$$

where

$$p_r = e_{(n, \lambda_0)} + (b_{1r} t_r + o(t_r)) e_{(n-1, \lambda_0)} + (b_{2r} t_r + o(t_r)) e_{(n-2, \lambda_0)} + o(t_r).$$

Let  $\lambda_{1r}, \dots, \lambda_{nr}$  denote the roots of the polynomials  $p_r$  for  $r = 1, 2, \dots$ , respectively. We have

$$(34) \quad \sum_{k=1}^n (\lambda_{kr} - \lambda_0) = -b_{1r} t_r + o(t_r),$$

and

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

Set  $z_{kr} = \lambda_{kr} - \lambda_0$  for  $k = 1, \dots, n$  and  $r = 1, 2, \dots$ . By (34) and (35) we have

$$(36) \qquad \qquad \qquad \sum_{k=1}^n z_{kr} = -b_{1r}t_r + o(t_r)$$

and

$$(37) \qquad \qquad \qquad \sum_{k=1}^n (z_{kr})^2 = -2b_{2r}t_r + o(t_r),$$

respectively. With this notation, (33) becomes

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

Taking second-order Taylor expansions yields

$$(38) \qquad \eta_r t_r + o(t_r) \geq \phi'(\lambda_0; z_{kr}) + \frac{1}{2} \phi''(\lambda_0; z_{kr}, z_{kr}) + o(|z_{kr}|^2), \quad k = 1, \dots, n,$$

(note that if  $\phi$  is quadratic then the term  $o(|z_{kr}|^2)$  equals zero). Since  $p_r \rightarrow e_{(n, \lambda_0)}$ , we have  $z_{kr} \rightarrow 0$ ,  $k = 1, \dots, n$ . 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  $\phi$  is quadratic then we can take  $\epsilon = 0$  without the assumption of positive definiteness). Therefore, for  $r \geq r_0$ , the inequalities (38) imply the inequalities

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

Now sum over  $k$  and again use the positive definiteness of  $\phi''(\lambda_0; \cdot, \cdot)$  with Lemma 2.5 and definition (10) 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 (36) and (37) gives the relation

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

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  $\phi$  is quadratic, we can obtain this inequality with  $\epsilon = 0$  without the positive definiteness assumption. Since  $\epsilon > 0$  was arbitrary, we obtain (32).  $\square$

If  $\phi$  is not quadratic and  $\phi''(\lambda_0; \cdot, \cdot)$  is only positive semidefinite, we can still sharpen (26) but not as finely as in (32). The proof in the indefinite case is completely different. Unlike the proof of Theorem (2.7), it relies only on the Gauss-Lucas Theorem.

**Theorem 2.8.** *If in Theorem 2.4 it is further assumed that  $\phi$  is twice continuously differentiable, then condition (26) can be strengthened to*

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

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

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

where

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

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

$$\begin{aligned} \mu &\geq \phi(\lambda_0) + \left\langle \phi'(\lambda_0), \frac{-a_1}{na_0} \right\rangle + \frac{1}{2} \left\langle \phi''(\lambda_0) \left( \frac{-a_1}{na_0} \right), \left( \frac{-a_1}{na_0} \right) \right\rangle \\ (40) \quad &+ \frac{1}{2} \langle \phi''(\lambda_0)r, r \rangle + o \left( \left| \frac{-a_1}{na_0} \pm r \right|^2 \right). \end{aligned}$$

Next, suppose that  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}((e_{(n, \lambda_0)}, \phi(\lambda_0)))$  with  $v$  given by (25). By Lemma 2.1,  $(\tilde{v}, \eta) \in T_{\text{epi}(\hat{\phi}_1)}(e_{(n, \lambda_0)}, \phi(\lambda_0))$ , where  $\tilde{v} = \sum_{k=1}^n b_k e_{(n-k, \lambda_0)}$ . Then, as in the proof of Theorem 2.4, there exist sequences  $t_j \downarrow 0$  and  $\{(p_j, \mu_j)\} \in \text{epi}(\hat{\phi}_1)$  such that

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

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

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

$t_j^{-1}(\mu_j - \phi(\lambda_0)) \rightarrow \eta$ ,  $a_0^j = 1$  for all  $j = 1, 2, \dots$ , and  $t_j^{-1}a_k^j \rightarrow b_k$  for  $k = 1, \dots, n$ . By replacing  $\mu$  by  $\mu_j$  and  $a_k$  by  $a_k^j$ ,  $k = 0, \dots, n$  in (40),

dividing through by  $t_j$ , and slightly re-arranging, we obtain

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

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  $j \rightarrow \infty$  yields (39).  $\square$

The representation (12) along with Theorems 2.4, 2.7, and 2.8 yield the following representations and bounds for the subderivative of the function  $\hat{\phi}$ .

**Theorem 2.9.** *Let  $\hat{\phi}$  be as defined in (1),  $\lambda_0 \in \text{dom}(\partial\phi)$  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 (27) and (28) are not satisfied; otherwise,*

$$(41) \quad d\hat{\phi}(e_{(n, \lambda_0)})(v) \geq \frac{1}{n} \phi'(\lambda_0; -b_1),$$

*with equality holding if  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$ .*

*If it is further assumed that the function  $\phi$  is twice continuously differentiable at  $\lambda_0$ , then whenever  $n > 1$  the inequality (41) can be refined to*

$$(42) \quad \frac{1}{n} \left[ \phi'(\lambda_0; -b_1) + \frac{1}{n-1} \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) \right] \\ \leq d\hat{\phi}(e_{(n, \lambda_0)})(v)$$

$$(43) \quad \leq \frac{1}{n} \left[ \phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) \right],$$



whenever (27) and (28) are both satisfied. Moreover, equality holds in (43) if any one of the following three conditions hold:

$$(44) \quad \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) = 0,$$

$$(45) \quad \phi''(\lambda_0; \cdot, \cdot) \text{ is positive definite, or}$$

$$(46) \quad \phi \text{ is quadratic.}$$

*Proof.* By Theorem 2.4 we know that the subderivative is  $+\infty$  if (27) and (28) are not satisfied. The lower bounds (41) and (42) are immediate consequences of Theorems 2.4 and 2.8, respectively.

The representation (12) and Lemma 2.1 imply that with no loss in generality we may assume for the remainder of the proof that  $b_0 = 0$  in  $v$ .

Suppose that  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$  and (27) and (28) hold. We show that equality must hold in (41). As noted in Theorem 2.4,  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$  implies that  $b_k = 0$ ,  $k = 2, 3, \dots, n$ . 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 (13) shows that

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

hence equality holds in (41).

If either  $\phi$  is quadratic, or  $\phi''(\lambda_0; \cdot, \cdot)$  is positive definite, then Theorem 2.7 tells us that the expression on the right hand side of (43) is also a lower bound. Thus, to establish equality in these two cases we need only establish the upper bound (43). In addition, once this upper bound is established then we also obtain equality when (44) holds since in this case the upper bound (43) reduces to the lower bound (41). Thus, it remains only to prove the upper bound (43). We assume throughout that the polynomial  $v$  satisfies both (27) and (28).

We use (13) to establish the upper bound (43). The bound is obtained by considering the tangents to smooth curves having as limit  $e_{(n, \lambda_0)}$ . The proof proceed by considering the even and odd cases for  $n$  separately. But in both cases we make use of the following family of

polynomials:

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

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  $\xi$ , this polynomial has only two roots:

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

For  $\xi$  real and positive, the second-order Taylor expansion of  $\phi$  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 (27),  $\langle \phi'(\lambda_0), \sqrt{-b_2} \rangle = 0$ . Therefore,

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

establishing the even case.

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

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

and consider the family of polynomials

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

For all values of  $\xi$  the roots of this polynomial are

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

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

$$(47) \quad \phi(\lambda_0) + (\xi/n) \left[ \phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) \right] + o(\xi).$$

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

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

also equals (47). Therefore, for  $\xi$  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.  $\square$

Theorem 2.4 and its refinements give necessary conditions for inclusion in the tangent cone  $T_{\text{epi}(\hat{\phi})}((e_{(n,\lambda_0)}, \phi(\lambda_0)))$ . We now use the conditions given in Theorem 2.9 to characterize the tangent cone when  $\phi$  is twice differentiable at  $\lambda_0$ .

**Theorem 2.10.** *Let  $\lambda_0 \in \text{dom}(\phi)$  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  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$  or  $\phi$  is twice continuously differentiable at  $\lambda_0$  and any one of the three conditions (44)-(46) hold, then  $(v, \eta) \in T_{\text{epi}(\hat{\phi})}((e_{(n,\lambda_0)}, \phi(\lambda_0)))$  if and only if*

$$(48) \quad \begin{aligned} \eta &\geq \frac{1}{n} \left[ \phi'(\lambda_0; -b_1) + \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}) \right], \\ 0 &= \left\langle g, \sqrt{-b_2} \right\rangle \quad \forall g \in \partial\phi(\lambda_0), \quad \text{and} \\ 0 &= b_k, \quad k = 3, \dots, n, \end{aligned}$$

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

*Proof.* Apply Theorem 2.9 in conjunction with the representation (12).  $\square$

## 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]:

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

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

$$(49) \quad f(y) \geq f(x) + \langle v, y - x \rangle + o(|y - x|).$$

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]

$$(50) \quad \widehat{N}_{\text{epi}(f)}(x) = \left\{ t(v, -1) \mid v \in \hat{\partial}f(x), t > 0 \right\} \cup \left\{ (v, 0) \mid v \in \hat{\partial}f(x)^\infty \right\},$$

where  $\hat{\partial}f(x)^\infty$  denotes the recession cone of the set  $\hat{\partial}f(x)$  [14, Definition 3.3]. The regular subdifferential is related to the subderivative by the formula [14, Exercise 8.4]

$$(51) \quad \hat{\partial}f(x) = \{v \mid \langle v, w \rangle \leq df(x)(w) \forall w\}.$$

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 representation (51) implies the inequality

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

We use the relation (51) 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. We begin by defining a parametrized family of multifunctions  $\Delta_\delta : \mathbb{C} \rightrightarrows \mathbb{C}^{n+1}$  with parameter values  $0 \leq \delta \in \mathbb{R}$ . For  $\delta = 0$ , set

$$\Delta_0(\lambda_0) = \left\{ \left( \begin{array}{c} \theta_0 \\ \vdots \\ \theta_n \end{array} \right) \mid \begin{array}{l} \theta_0 = 0, \theta_1 \in -\frac{1}{n}\partial\phi(\lambda_0), \\ \langle \theta_2, g^2 \rangle \leq 0 \quad \forall g \in \partial\phi(\lambda_0) \end{array} \right\}.$$

For  $\delta > 0$ , the multifunction  $\Delta_\delta$  is only defined when  $\phi$  is twice continuously differentiable in which case  $(\theta_0, \theta_1, \dots, \theta_n)^T \in \Delta_\delta(\lambda_0)$  if and only if

$$(52) \quad \theta_0 = 0, \quad \theta_1 = \frac{-1}{n} \phi'(\lambda_0),$$

and  $\theta_2$  satisfies

$$(53) \quad \langle \theta_2, \phi'(\lambda_0)^2 \rangle \leq \delta \langle (i \phi'(\lambda_0)), \phi''(\lambda_0)(i \phi'(\lambda_0)) \rangle.$$

Given  $\lambda_0 \in \text{dom}(\partial\phi)$  and  $\delta \geq 0$ , the set  $\Delta_\delta(\lambda_0)$ , when defined, is a non-empty closed convex set. The recession cone of  $\Delta_\delta(\lambda_0)$  is independent of the choice of  $\delta \geq 0$ :

$$(54) \quad \begin{aligned} \Delta_\delta(\lambda_0)^\infty &= \left\{ \begin{pmatrix} \theta_0 \\ \vdots \\ \theta_n \end{pmatrix} \mid \begin{array}{l} \theta_0 = 0, \theta_1 = 0, \\ \langle \theta_2, g^2 \rangle \leq 0 \quad \forall g \in \partial\phi(\lambda_0) \end{array} \right\} \\ &= \{0\}^2 \times (\mathbb{R}_+ \partial\phi(\lambda_0))^\circ \times \mathbb{C}^{n-2}, \end{aligned}$$

where  $\mathbb{R}_+ \partial\phi(\lambda_0) = \{\mu\omega \mid \mu \in \mathbb{R}_+, \omega \in \partial\phi(\lambda_0)\}$ . We now characterize the support function for the sets  $\Delta_\delta(\lambda_0)$ .

**Lemma 3.1.** *Let  $\lambda_0 \in \text{dom}(\phi)$  be such that  $\partial\phi(\lambda_0) \neq \{0\}$ , and let  $\delta \geq 0$ . Then for every vector*

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

*we have  $\sigma_{\Delta_\delta(\lambda_0)}(b) = +\infty$  if the components of  $b$  do not satisfy (27) and (28); otherwise,*

$$(55) \quad \sigma_{\Delta_\delta(\lambda_0)}(b) = \frac{1}{n} \phi'(\lambda_0; -b_1) + \delta \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2}).$$

*The term  $\delta \phi''(\lambda_0; \sqrt{-b_2}, \sqrt{-b_2})$  is to be interpreted as zero with  $\delta = 0$  whenever  $\phi$  is not twice continuously differentiable at  $\lambda_0$ , and equals zero if  $\phi$  is twice continuously differentiable at  $\lambda_0$  with*

$$\langle (i \phi'(\lambda_0)), \phi''(\lambda_0)(i \phi'(\lambda_0)) \rangle = 0.$$

*Proof.* Set

$$\Delta_{(\delta, 2)}(\lambda_0) = \{\theta_2 \mid \theta \in \Delta_\delta(\lambda_0)\}.$$

Then

$$\begin{aligned}
\sigma_{\Delta_\delta(\lambda_0)}(b) &= \sup_{\theta \in \Delta_\delta(\lambda_0)} \langle \theta, b \rangle \\
&= \sup \left\{ \sum_{j=1}^n \langle \theta_j, b_j \rangle \left| \begin{array}{l} \theta_1 \in -\frac{1}{n} \partial \phi(\lambda_0) \\ \theta_2 \in \Delta_{(\delta, 2)}(\lambda_0) \\ \theta_j \in \mathbb{C}, \quad j = 3, \dots, n \end{array} \right. \right\} \\
(56) \quad &= \frac{1}{n} \phi'(\lambda_0; -b_1) + \sup_{\theta_2 \in \Delta_{(\delta, 2)}(\lambda_0)} \langle \theta_2, b_2 \rangle + \sum_{j=3}^n \sup_{\theta_j \in \mathbb{C}} \langle \theta_j, b_j \rangle,
\end{aligned}$$

where the second equality follows the definition of  $\Delta_\delta(\lambda_0)$  and the third equality follows from (18).

Observe that the final term in (56) implies that  $\sigma_{\Delta_\delta(\lambda_0)}(b) = +\infty$  if  $b$  does not satisfy (28). We now show that (27) must also be satisfied. Let  $g$  be a nonzero subgradient of  $\phi$  at  $\lambda_0$  and let  $\theta_2 \in \Delta_{(\delta, 2)}(\lambda_0)$ . Since  $g$  is nonzero, there exist real scalars  $\mu_1$  and  $\mu_2$  such that

$$(57) \quad \theta_2 = \mu_1 g^2 + \mu_2 i g^2,$$

and so

$$(58) \quad \langle \theta_2, g^2 \rangle = \mu_1 |g|^4.$$

In particular, this implies that the inequalities  $\langle \theta_2, g^2 \rangle \leq 0$  when  $\delta = 0$  and (53) when  $\delta > 0$  do not restrict the values that  $\mu_2$  may take while  $\mu_1$  is allowed take arbitrarily large negative values. Now, as with  $\theta_2$ , there exist real scalars  $\beta_1$  and  $\beta_2$  such that

$$b_2 = \beta_1 g^2 + \beta_2 i g^2$$

so that

$$\langle \theta_2, b_2 \rangle = \mu_1 \beta_1 |g|^4 + \mu_2 \beta_2 |g|^4.$$

The fact that  $\mu_2$  is unrestricted implies that the second term in (56) is  $+\infty$  unless  $\beta_2 = 0$ . Similarly, the fact that  $\mu_1$  can take arbitrarily large negative values implies that the second term in (56)  $+\infty$  unless  $\beta_1 \geq 0$ . Therefore,

$$(59) \quad \sigma_{\Delta_\delta(\lambda_0)}(b) < \infty \text{ only if } b_2 = \beta_1 g^2 \text{ with } \beta_1 \geq 0,$$

or equivalently,

$$0 = \left\langle g, \sqrt{-b_2} \right\rangle.$$

Since  $g$  is an arbitrary nonzero element of  $\partial \phi(\lambda_0)$ , we have that (27) holds. That is,  $\sigma_{\Delta_\delta(\lambda_0)}(b)$  is finite if and only if (27) and (28) are satisfied.

Next assume that  $b$  satisfies (27) and (28) and consider the cases  $\delta = 0$  and  $\delta > 0$  separately. If  $\delta = 0$ , then, by (57), (58), and (59), the second term in (56) takes the form

$$\sup_{\mu_1 \leq 0} \mu_1 \beta_1 |g|^4 ,$$

where  $\beta_1 \geq 0$ . This supremum is obviously zero, thus establishing (55). On the other hand, if  $\delta > 0$ , then, again by (57), (58), and (59), the second term in (56) is the supremum of the values  $\mu_1 \beta_1 |\phi'(\lambda_0)|^4$  over all values of  $\mu_1$  satisfying

$$\mu_1 |\phi'(\lambda_0)|^4 \leq \delta \langle (i \phi'(\lambda_0)) , \phi''(\lambda_0) (i \phi'(\lambda_0)) \rangle .$$

Again, since  $\beta_1 \geq 0$ , this supremum is given by

$$\begin{aligned} & \delta \beta_1 \langle (i \phi'(\lambda_0)) , \phi''(\lambda_0) (i \phi'(\lambda_0)) \rangle \\ &= \delta \left\langle \left( i \sqrt{\beta_1} \phi'(\lambda_0) \right) , \phi''(\lambda_0) \left( i \sqrt{\beta_1} \phi'(\lambda_0) \right) \right\rangle \\ &= \delta \left\langle \left( \sqrt{-b_2} \right) , \phi''(\lambda_0) \left( \sqrt{-b_2} \right) \right\rangle , \end{aligned}$$

where the second equality follows from (59). Hence, (55) is valid in this case as well and the final statement of the lemma also holds.  $\square$

Lemma 3.1 combined with Theorem 2.9 and the representation (51) 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  $\tau_{\lambda_0} : \mathcal{P}^n \rightarrow \mathbb{C}^{n+1}$  to be the mapping that takes a polynomial to its Taylor series coefficients when expanded at the base point  $\lambda_0$ , specifically,

$$(60) \quad \tau_{\lambda_0}(p) = (p(\lambda_0), p'(\lambda_0), \dots, p^{(n)}(\lambda_0)/n!)^T .$$

Equivalently, if  $p$  has the representation  $p = \sum_{k=0}^n a_k e_{(n-k, \lambda_0)}$ , then  $\tau_{\lambda_0}(p) = (a_n, a_{n-1}, \dots, a_0)^T$ . The family of linear transformations  $\tau_\lambda$  is continuous in  $\lambda$ , and for each  $\lambda_0$  the transformation  $\tau_{\lambda_0}$  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 3.2.** *Let  $\hat{\phi}$  be as defined in (1) and  $\lambda_0 \in \text{dom}(\phi)$  be such that  $\partial\phi(\lambda_0) \neq \{0\}$ . Then*

$$(61) \quad d\hat{\phi}(e_{(n, \lambda_0)})(v) \geq \sigma_{\Delta_0(\lambda_0)}(\tau_{\lambda_0}(v))$$

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

$$(62) \quad \tau_{\lambda_0}^* \Delta_0(\lambda_0) \subset \hat{\partial}\hat{\phi}(e_{(n, \lambda_0)}) ,$$

where  $\tau_{\lambda_0}^*$  is the adjoint of  $\tau_{\lambda_0}$  with respect to the inner product  $\langle \cdot, \cdot \rangle_{\lambda_0}$ . Equality holds in both (61) and (62) if  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$ .

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

$$(63) \quad \sigma_{\Delta_{\delta_1}(\lambda_0)}(\tau_{\lambda_0}(v)) \leq d\hat{\phi}(e_{(n,\lambda_0)})(v) \leq \sigma_{\Delta_{\delta_2}(\lambda_0)}(\tau_{\lambda_0}(v)),$$

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

$$(64) \quad \tau_{\lambda_0}^* \Delta_{\delta_1}(\lambda_0) \subset \hat{\partial}\hat{\phi}(e_{(n,\lambda_0)}) \subset \tau_{\lambda_0}^* \Delta_{\delta_2}(\lambda_0)$$

where  $\delta_1 = 1/(n(n-1))$  and  $\delta_2 = 1/n$ . Furthermore, if  $\phi$  is quadratic, or  $\phi''(\lambda_0; \cdot, \cdot)$  is positive definite, or  $\langle (i\phi'(\lambda_0)), \phi''(\lambda_0)(i\phi'(\lambda_0)) \rangle = 0$ , 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 (61) follows from the bound (41) in Theorem 2.9 coupled with Lemma 3.1. The left and right inequalities in (63) follow from (42) and (43) in Theorem 2.9, respectively, again in conjunction with Lemma 3.1. The subdifferential inclusions in (62) and the left hand side of (64) follow immediately from the definition of the adjoint transformation, the representation (51) and the inequalities in (61) and the left hand side of (63), 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 (64) 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 2.9, the final statement of Lemma 3.1, and the preceding comment on the relationship between support functions and convex sets.  $\square$

We now apply Theorem 3.2 to (50) obtaining approximations to the cone of regular normals to  $\text{epi}(\hat{\phi})$  at the point  $(e_{(n,\lambda_0)}, \phi(\lambda_0))$ . We begin by extending the definitions for the sets  $\Delta_\delta(\lambda_0)$ . For  $\delta = 0$  define

$$\begin{aligned} \Xi_0(\lambda_0) &= \{\gamma(\theta, -1) \mid \theta \in \Delta_0(\lambda_0), 0 < \gamma \in \mathbb{R}\} \cup \Delta_0(\lambda_0)^\infty \\ &= \left\{ (w, \mu) \in \mathbb{C}^{n+1} \times \mathbb{R} \mid \begin{array}{l} \mu \leq 0, w_0 = 0, w_1 \in \frac{\mu}{n} \partial\phi(\lambda_0), \\ \text{and } \langle w_2, g^2 \rangle \leq 0 \forall g \in \partial\phi(\lambda_0) \end{array} \right\}. \end{aligned}$$



For  $0 < \delta \in \mathbb{R}$ , we assume that  $\phi$  is twice continuously differentiable at  $\lambda_0$  and define

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

so that  $(w, \mu) \in \Xi_\delta(\lambda_0)$  if and only if

$$\begin{aligned} \mu \leq 0, \quad w_0 = 0, \quad w_1 = \frac{\mu}{n} \phi'(\lambda_0), \quad \text{and} \\ \langle w_2, \phi'(\lambda_0)^2 \rangle \leq -\mu \delta \langle (i \phi'(\lambda_0)), \phi''(\lambda_0)(i \phi'(\lambda_0)) \rangle, \end{aligned}$$

where  $w = (w_0, w_1, \dots, w_n)^T \in \mathbb{C}^{n+1}$  and  $\mu \in \mathbb{R}$ . Also, for each  $\lambda \in \mathbb{C}$ , define the linear transformation  $\hat{\tau}_\lambda : \mathcal{P}^n \times \mathbb{R} \rightarrow \mathbb{C}^{n+1} \times \mathbb{R}$  by  $\hat{\tau}_\lambda(p, \mu) = (\tau_\lambda(p), \mu)$  where the linear transformation  $\tau_\lambda$  is defined in (60). 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  $\mathcal{P}^n \times \mathbb{R}$ , is the linear transformation  $\hat{\tau}_\lambda^* : \mathbb{C}^{n+1} \times \mathbb{R} \rightarrow \mathcal{P}^n \times \mathbb{R}$  given by

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

Theorem 3.2 and relation (50) give the following corollary to Theorem 3.2.

**Corollary 3.3.** *Let  $\delta \geq 0$  and  $\hat{\phi}$  be as defined in (1) with  $\lambda_0 \in \text{dom}(\partial\phi)$  satisfying  $\partial\phi(\lambda_0) \neq \{0\}$ . Then*

$$(65) \quad \hat{\tau}_{\lambda_0}^* \Xi_0(\lambda_0) \subset \hat{N}_{\text{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0)),$$

with equality holding if  $\text{int}(\partial\phi(\lambda_0)) \neq \emptyset$ . If it is further assumed that the function  $\phi$  is twice continuously differentiable at  $\lambda_0$  with  $\phi'(\lambda_0) \neq 0$ , then

$$(66) \quad \hat{\tau}_{\lambda_0}^* \Xi_{\delta_1}(\lambda_0) \subset \hat{N}_{\text{epi}(\hat{\phi})}(e_{(n, \lambda_0)}, \phi(\lambda_0)) \subset \hat{\tau}_{\lambda_0}^* \Xi_{\delta_2}(\lambda_0)$$

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

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

#### 4. THE ABCISSA MAPPING

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

$$\mathbf{a}(p) = \sup \{\text{Re } \lambda \mid \lambda \in \mathcal{R}(p)\}.$$

Here  $\mathbf{a} = \hat{\phi}$ , where  $\hat{\phi}$  is defined in (1) with the function  $\phi$  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 4.1.** *Given  $\lambda_0 \in \mathbb{C}$ , one has  $(v, \eta) \in T_{\text{epi}(\mathbf{a})}(e_{(n, \lambda_0)}, \text{Re}(\lambda_0))$ , with*

$$(67) \quad v = \sum_{k=0}^n b_k e_{(n-k, \lambda_0)}$$

*if and only if*

$$(68) \quad \eta \geq -\frac{1}{n} \text{Re}(b_1),$$

$$(69) \quad 0 \leq \text{Re } b_2, \quad 0 = \text{Im } b_2, \quad \text{and}$$

$$(70) \quad 0 = b_k, \quad k = 3, \dots, n.$$

*Moreover,  $d\mathbf{a}(e_{(n, \lambda_0)})(v) = +\infty$  if (69) and (70) are not satisfied; otherwise,*

$$(71) \quad d\mathbf{a}(e_{(n, \lambda_0)})(v) = -\frac{1}{n} \text{Re}(b_1).$$

*Proof.* The final statement of the theorem follows immediately from the final statement of Theorem 2.9. The first part of the result follows from Theorem 2.10.  $\square$

**Remark** The two conditions  $0 \leq \text{Re } b_2$  and  $0 = \text{Im } b_2$  in (69) are equivalent to the single condition  $0 = \text{Re } \sqrt{-b_2}$  which follows from condition (48) in Theorem 2.10.

The subdifferential and normal cone characterizations follow directly from Theorem 3.2 and Corollary 3.3.

**Theorem 4.2.** *Let  $\lambda_0 \in \mathbb{C}$  be given. Then*

$$\widehat{N}_{\text{epi}(\mathbf{a})}(e_{(n, \lambda_0)}, \text{Re}(\lambda_0)) = \hat{\tau}_{\lambda_0}^* \left\{ t(w, -1) \mid \begin{array}{l} 0 \leq t, w_0 = 0, w_1 = \frac{-1}{n} \\ \text{and } \text{Re}(w_2) \leq 0 \end{array} \right\},$$

*and*

$$\partial\mathbf{a}(e_{(n, \lambda_0)}) = \tau_{\lambda_0}^* \{w \mid w_0 = 0, w_1 = -1/n, \text{ and } \text{Re}(w_2) \leq 0\}.$$

The results of Theorems 4.1 and 4.2 coincide precisely with those found in [3, 5].

## 5. THE RADIUS MAPPING

Consider now the radius mapping for polynomials:

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

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

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

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

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

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

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

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 5.1.** *Let  $p \in \mathcal{M}^n$  be any polynomial for which  $r(p) > 0$ . Then  $(v, \eta) \in T_{\text{epi}(r)}(p, \mu)$  if and only if  $(v, \mu\eta) \in T_{\text{epi}(r_2)}(p, \frac{1}{2}\mu^2)$ , where*

$$r_2(p) = \sup \left\{ \frac{1}{2} |\lambda|^2 \mid \lambda \in \mathcal{R}(p) \right\}.$$

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

$$(72) \quad \{(p_k, \mu_k)\} \subset \text{epi}(r) \text{ and } t_k \searrow 0$$

such that

$$(73) \quad \frac{p_k - p}{t_k} \rightarrow v, \text{ and}$$

$$(74) \quad \frac{\mu_k - \mu}{t_k} \rightarrow \eta.$$

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

Now since  $(p, \mu) \in \text{epi}(r)$  if and only if  $(p, \mu^2/2) \in \text{epi}(r_2)$ , we have (72) is equivalent to

$$(75) \quad \{(p_k, \mu_k^2/2)\} \subset \text{epi}(r_2) \text{ and } t_k \searrow 0.$$

Also, since  $0 < \mu, \mu_k$ , (74) is equivalent to

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

or equivalently,

$$(76) \quad \frac{\frac{1}{2}\mu_k^2 - \frac{1}{2}\mu^2}{t_k} \rightarrow \mu\eta.$$

Therefore, the statements (72)–(74) are equivalent to the statements (75), (73), and (76), or equivalently,  $(v, \mu\eta) \in T_{\text{epi}(r_2)}(p, \mu)$ .  $\square$

Lemma 5.1 gives the representation

$$(77) \quad T_{\text{epi}(r)}(p, \mu) = \{(v, \eta/\mu) \mid (v, \eta) \in T_{\text{epi}(r_2)}(p, \mu^2/2)\},$$

whenever  $r(p) > 0$ . Since  $\frac{1}{2}|\cdot|^2$  is quadratic, with  $(\frac{1}{2}|\cdot|^2)''(\zeta; \delta, \delta) = |\delta|^2$ , Theorem 2.10 provides a complete characterization of the tangent cone  $T_{\text{epi}(r)}(e_{(n, \lambda_0)}, |\lambda_0|)$ .

**Theorem 5.2.** *Let  $\lambda_0 \in \mathbb{C}$  and let  $(v, \eta) \in \mathcal{P}^n \times \mathbb{R}$  be such that*

$$(78) \quad v = \sum_{k=0}^n b_k e_{(n-k, \lambda_0)}.$$

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

$$(79) \quad \eta \geq \frac{1}{n} |b_1|,$$

$$(80) \quad 0 = b_k, \quad k = 2, 3, \dots, n.$$

*Moreover,  $dr(e_{(n, \lambda_0)})(v) = +\infty$  if (80) is not satisfied; otherwise,*

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

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

$$(81) \quad \eta \geq \frac{1}{n |\lambda_0|} [|b_2| - \text{Re } \bar{\lambda}_0 b_1],$$

$$(82) \quad 0 = \text{Re } \bar{\lambda}_0 \sqrt{-b_2}, \text{ and}$$

$$(83) \quad 0 = b_k, \quad k = 3, \dots, n.$$

*Moreover,  $dr(e_{(n, \lambda_0)})(v) = +\infty$  if (82) and (83) are not satisfied; otherwise,*

$$dr(e_{(n, \lambda_0)})(v) = \frac{1}{n |\lambda_0|} [|b_2| - \text{Re } \bar{\lambda}_0 b_1].$$

*Proof.* The case  $\lambda_0 = 0$  follows from Theorem 2.9 and the representation (12) since  $\mathbf{B} = \partial |\cdot| (0)$  has non-empty interior, while the case  $\lambda_0 \neq 0$  follows from Theorem 2.10 and the representation (77).  $\square$

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

**Theorem 5.3.** *Let  $\lambda_0 \in \mathbb{C}$  and let the linear transformation  $\tau_\lambda : \mathcal{P}^n \rightarrow \mathbb{C}^{n+1}$  be as defined in (60). Set*

$$\Delta^r(\lambda_0) = \left\{ \left( \begin{array}{c} \theta_0 \\ \vdots \\ \theta_n \end{array} \right) \mid \theta_0 = 0, \theta_1 \in \frac{1}{n}\mathbf{B} \right\},$$

if  $\lambda_0 = 0$ ; otherwise, set

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

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

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

and

$$N_{\text{epi}(r)}(e_{(n,\lambda_0)}, 0) = \left\{ (\tau_{\lambda_0}^*(w), -\mu) \mid \begin{array}{l} \mu \geq 0, w \in \mathbb{C}^{n+1}, \\ w_0 = 0, |w_1| \leq \mu \end{array} \right\}.$$

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

$$(84) \quad dr(e_{(n,\lambda_0)})(v) = \sigma_{\Delta^r(\lambda_0)}(\tau_{\lambda_0}(v)),$$

$$(85) \quad \hat{d}r(e_{(n,\lambda_0)}) = \tau_{\lambda_0}^* \Delta^r(\lambda_0),$$

and

$$(86) \quad N_{\text{epi}(r)}(e_{(n,\lambda_0)}, \lambda_0) = \left\{ (\tau_{\lambda_0}^*(w), -\mu) \mid \begin{array}{l} \mu \geq 0, w_0 = 0, w_1 = \frac{-\mu\lambda_0}{|\lambda_0|}, \\ \langle w_2, \lambda_0^2 \rangle \leq \mu |\lambda_0| \end{array} \right\}.$$

*Proof.* Let us first suppose that  $\lambda_0 = 0$ . In this case, the subdifferential  $\mathbf{B} = \partial |\cdot| (0)$  has non-empty interior and so the results of Theorems 2.9 and (66) directly apply to give the result.

Next suppose that  $\lambda_0 \neq 0$ . In this case, Lemma 3.1 combined with Part (2) of Theorem 5.2 gives (84). This in turn establishes (85) due to the relation (51). The final relation (86) follows from the equivalence (50).  $\square$

## 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 2.4). However, substantial additional work was required to obtain the second-order necessary and sufficient conditions given in Theorem 2.10. It is gratifying that the second-order result preserves the simplicity and geometric appeal of Theorem 2.4. 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 squareroot splitting of the roots. This is nicely illustrated in the application to the radius mapping in Section 5.

However, Theorem 2.10 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  $\phi$  is twice differentiable. If true, this result would make a number of results possible including the extension of Theorem 2.10 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 the verification of subdifferential regularity remain to be established and will appear in a subsequent paper.

## REFERENCES

- [1] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimizing matrix stability. *Proc. of the American Math. Soc.*, 129:1635–1642, 2000.
- [2] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimal stability and eigenvalue multiplicity. *Foundations of Computational Mathematics*, 1:205–225, 2001.
- [3] J.V. Burke, A.S. Lewis, and M.L. Overton. Variational analysis of the abscissa mapping for polynomials via the Gauss-Lucas theorem. *Journal of Global Optimization*, 28:259–268, 2004.
- [4] J.V. Burke and M.L. Overton. Variational analysis of the abscissa mapping for polynomials. *SIAM Journal on Control and Optimization*, 39:1651–1676, 2000.
- [5] J.V. Burke and M.L. Overton. Variational analysis of non-Lipschitz spectral functions. *Math. Programming*, 90:317–351, 2001.
- [6] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York 1983. Republished by SIAM, 1990.
- [7] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998.
- [8] L.V. Levantovskii. On singularities of the boundary of a stability region. *Moscow Univ. Math. Bull.*, 35:19–22, 1980.

- [9] F. Lucas. Sur une application de la mécanique rationnelle à la théorie des équations. *C.R. Acad. Sci. Paris*, 89:224—226, 1879.
- [10] F. Lucas. Géométrie des Polynômes. *J. École Polytech*, 46:1—33, 1879.
- [11] M. Marden. *Geometry of Polynomials*. American Mathematical Society, 1966.
- [12] B.S. Mordukhovich. *Approximation Methods in Problems of Optimization and Control*. Nauka, Moscow, 1988. Republished by Wiley, 2001.
- [13] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [14] R.T. Rockafellar and R.J.B. Wets. *Variational Analysis*. Springer, 1998.

1. DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195.

*E-mail address:* burke@math.washington.edu

2. DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY, BURNABY, BRITISH COLUMBIA, CANADA V5A 1S6.

*E-mail address:* aslewis@sfu.ca

3. COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK, NY 10012.

*E-mail address:* overton@cs.nyu.edu