Nonlinear Analysis 75 (2012) 1168–1187

Contents lists available at SciVerse ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

On the subdifferential regularity of max root functions for polynomials

James V. Burke [a], Julia Eaton [b,*]

[a] *Department of Mathematics, University of Washington, Seattle, WA, United States*

[b] *Department of Mathematics and Computer Science, University of Puget Sound, Tacoma, WA, United States*

ARTICLE INFO

Keywords:
Spectral radius
Spectral abscissa
Polynomial root functions
Subdifferential regularity
Max polynomial root functions

ABSTRACT

In 2001, Burke and Overton showed that the abscissa mapping on polynomials is subdifferentially regular on the monic polynomials of degree n. We extend this result to the class of max polynomial root functions which includes both the polynomial abscissa and the polynomial radius mappings. The approach to the computation of the subgradient simplifies that given by Burke and Overton and provides new insight into the variational properties of these functions.

1. Introduction

Let $\mathcal{P}^n$ denote the linear space of polynomials over $\mathbb{C}$ of degree n or less. The abscissa and radius mappings on $\mathcal{P}^n$ are given by

$$\mathbf{a}(p) = \max\{\mathrm{Re}(\lambda) \mid \lambda \in \mathcal{R}(p)\} \quad \text{and} \quad \mathbf{r}(p) = \max\{|\lambda| \mid \lambda \in \mathcal{R}(p)\},$$

respectively, where $\mathcal{R}(p) = \{\lambda \mid p(\lambda) = 0\}$. When composed with the characteristic polynomial of the $n \times n$ matrix A, the resulting mappings are called the spectral abscissa and the spectral radius, respectively. These mappings characterize the asymptotic stability of solutions to linear dynamical systems, and so understanding their variational properties assists in understanding the variational behavior of stability [1]. In [2], Burke and Overton use techniques from variational analysis [3–5] to give formulas for the subdifferential of the abscissa mapping $\mathbf{a}$ and establish its subdifferential regularity on the affine set of monic polynomials of degree n. The proof of subdifferential regularity has three challenging steps. The first uses a technique developed by Levantovskii in [6] to characterize the tangent cone to the epigraph of the abscissa mapping $\mathbf{a}$ at the polynomial $(\lambda - \lambda_0)^n$. This step requires several pages of dense computation. The tangent cone representation is then used to provide a formula for the subderivative of $\mathbf{a}$ at $(\lambda - \lambda_0)^n$. In the second step, the set of regular normals to the epigraph is computed for a general monic polynomial. The representation for the regular normal cone yields a formula for the regular subdifferential of $\mathbf{a}$ on the monics. In the third and final step, the set of limiting regular normals is computed and subdifferential regularity is established.

In [7] it is shown that the Gauss–Lucas Theorem [8] can be applied to dramatically simplify the first step; the computation of the tangent cone to the epigraph of the abscissa mapping $\mathbf{a}$ at $(\lambda - \lambda_0)^n$. The Gauss–Lucas technique is used in [9] to extend these variational results to a much broader class of functions of the roots of polynomials which we call *max polynomial root functions*. This class includes both the abscissa and radius mappings, and so opens the door to a deeper understanding of the variational behavior of a large class of important functions of the roots of polynomials. However, the results in [9] do not address steps 2 and 3 of [2]. Following the work in [10], these steps are addressed here with the goal of extending the results of [2] to the class of max polynomial root functions. Although we rely on the underlying *factorization space* structure developed in [2], our approach differs significantly since we do not use epigraphical normal cones to compute subgradients.

* Corresponding author.
E-mail addresses: burke@math.washington.edu (J.V. Burke), jeaton@pugetsound.edu (J. Eaton).

0362-546X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.na.2011.01.021

Rather we go directly from the subderivative to the regular subdifferential and then on to the (limiting) subdifferential, short circuiting the normal cone computations. The key to the new derivation is to first establish, and then exploit, the sublinearity of the subderivative. This allows us to directly compute the regular subdifferential using well-known properties of support functionals.

We begin in Section 2 by recalling and slightly refining the notation used in [2,9]. The classes of functions under investigation are precisely defined, factorization spaces are introduced, and extensions are given to the basic results in [2] concerning the epigraphical tangent cones. In Section 3, we use the tangent cone results to develop formulas for the subderivatives. These results differ from those of [2,9] since we also show that the subderivatives are sublinear. This key difference sets the stage for a direct and simplified derivation of the regular subdifferential using elementary properties of support functions. In Section 4, we prepare for the subdifferential analysis by building inner products compatible with the linear mappings between $\mathcal{P}^n$, the factorization spaces, and $\mathbb{C}^{n+1}$. With these Euclidean structures in place, we derive formulas for the regular subdifferential in Section 5. Finally, in Section 6 we establish the subdifferential regularity of max polynomial root functions generated by convex functions that are either quadratic or whose Hessian is positive definite at all active roots. This result is used in Section 7 to establish the subdifferential regularity of the radius mapping on the set of monic polynomials of degree n.

As noted above, we use the methods of variational analysis as developed in [3–5]. To assist the reader we catalog some of the key tools and notation from these references. Let E be a Euclidean space, i.e. a finite dimensional real inner product space. In this paper, the scalar field is always $\mathbb{C}$, and all inner products can be represented as the real part of a Hermitian inner product on the underlying Euclidean space E. Let C be a nonempty subset of E. The *tangent cone* to C at a point $x \in C$ is given by

$$T_C(x) = \{d \mid \exists\{x^\nu\} \subset C, \{t_\nu\} \subset \mathbb{R}_+ \text{ such that } x^\nu \to x,\ t_\nu \downarrow 0 \text{ and } t_\nu^{-1}(x^\nu - x) \to d\}.$$

The tangent cone is a closed subset of E [5, Proposition 6.2]. A tangent vector $d \in T_C(x)$ is said to be *derivable* if there is a trajectory $\gamma : [0, \varepsilon] \to C$ with $\varepsilon > 0$ such that $\gamma(0) = x$ and $\gamma'(0) = d$. The set C is said to be *geometrically derivable* at x if every tangent direction to C at x is derivable. The *polar* of C is the set $C^\circ = \{w \mid \langle w, v\rangle \le 1 \text{ for all } v \in C\}$. The set C° is always closed and convex, and if C is closed and convex, then $(C^\circ)^\circ = C$. In general, $(C^\circ)^\circ$ is the closed convex hull of C, $\mathrm{cl}(\mathrm{con}(C))$. If C is a cone, then $C^\circ = \{w \mid \langle w, v\rangle \le 0 \text{ for all } v \in C\}$. The *regular normal cone* to a point $x \in C$ is the set $\widehat{N}_C(x) = (T_C(x))^\circ = \{z \mid \langle z, v\rangle \le 0 \text{ for all } v \in T_C(x)\}$. The *horizon cone* of C (also known as the *asymptotic cone* [11,12]) is the set

$$C^\infty = \{z \in E \mid \exists\{x^\nu\} \subset C, \{t_\nu\} \subset \mathbb{R}_+ \text{ s.t. } t_\nu \downarrow 0 \text{ and } t_\nu x^\nu \to z\}.$$

The horizon cone is always a closed cone. If C is convex, it can be shown that C^∞ is the usual *recession cone* from convex analysis. The *support function* of C is given by

$$\sigma_C(v) = \sup_{z \in C} \langle z, v\rangle .$$

A function is said to be *proper* if there is a point in its domain space where it takes a finite value. By [5, Theorem 8.24], there is a one-to-one correspondence between sublinear, lower semi-continuous (lsc) proper functions φ and nonempty, closed, convex subsets C of E such that $\sigma_C(x) = \varphi(x)$ for all $x \in E$.

Let $h : E \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. The *essential domain* of h is $\mathrm{dom}(h) = \{x \in E \mid h(x) < \infty\}$. In particular, h is proper if its essential domain is nonempty. The *epigraph* of h is given by $\mathrm{epi}(h) = \{(x, \beta) \in E \times \mathbb{R} \mid h(x) \le \beta\}$. The *subderivative* of h is the map $\mathrm{d}h(x) : E \to \overline{\mathbb{R}}$ given by

$$\mathrm{d}h(x)(\bar{v}) = \liminf_{t \downarrow 0, v \to \bar{v}} \frac{h(x + tv) - h(x)}{t}.$$

It generalizes the notion of directional derivative to nondifferentiable functions. The tangent cone to the epigraph and the subderivative are related by the formula

$$\mathrm{epi}(\mathrm{d}h(x)) = T_{\mathrm{epi}(h)}(x, h(x)) \tag{1}$$

[5, Theorem 8.2]. In particular, $\mathrm{d}h(x)(v) = \inf\{\eta \mid (v, \eta) \in T_{\mathrm{epi}(h)}(x, h(x))\}$. The *regular subdifferential* of h at $x \in \mathrm{dom}(h)$ is the set of *regular subgradients*:

$$\hat{\partial} h(x) = \{v \mid h(y) \ge h(x) + \langle v, y - x\rangle + o(\|y - x\|) \quad \forall y \in E\}.$$

The regular subdifferential is always a closed and convex subset of E. The subderivative and regular subdifferential are related by $\hat{\partial} h(x) = \{z \mid \langle z, v\rangle \le \mathrm{d}h(x)(v)\ \forall v \in E\}$ [5, Exercise 8.4]. In particular, we have $\sigma_{\hat{\partial} h(x)}(v) \le \mathrm{d}h(x)(v)$ for all $v \in E$. Therefore,

$$\text{if } \hat{\partial} h(x) \ne \emptyset \quad \text{and } \mathrm{d}h(x) \text{ is sublinear, lsc and proper, then } \sigma_{\hat{\partial} h(x)} = \mathrm{d}h(x). \tag{2}$$

We can obtain the regular normal cone to a point in $\mathrm{epi}(h)$ from the regular subdifferential and vise versa by the relationships below [5, Theorem 8.9]:

$$\widehat{N}_{\mathrm{epi}(h)}(x, h(x)) = \{t(z, -1) \mid z \in \hat{\partial} h(x), t > 0\} \cup \{(z, 0) \mid z \in \hat{\partial} h(x)^\infty\}, \tag{3}$$

and $\hat{\partial} h(x) = \{z \mid (z, -1) \in \widehat{N}_{\mathrm{epi}(h)}(x, h(x))\}$. If h is proper and convex, then the subderivative reduces to the usual notion of directional derivative, $h'(x; \cdot) = \mathrm{d}h(x)(\cdot) = \sigma_{\partial h(x)}(\cdot)$ with the regular subgradients corresponding to the usual subgradients of convex analysis. The general subdifferential of h at x is given by

$$\partial h(x) = \left\{ z \,\middle|\, \begin{array}{l} \exists x^\nu \to x,\ x^\nu \in \mathrm{dom}(h), \\ \exists z^\nu \in \hat{\partial} h(x^\nu) \text{ with } h(x^\nu) \to h(x) \text{ and } z^\nu \to z \end{array} \right\},$$

and the horizon subdifferential to h at x is given by

$$\partial^\infty h(x) = \left\{ z \,\middle|\, \begin{array}{l} \exists\, x^\nu \to x, \\ z^\nu \in \hat{\partial} h(x^\nu),\ \beta_\nu \downarrow 0 \\ \text{with } h(x^\nu) \to h(x) \text{ and } \beta_\nu z^\nu \to z \end{array} \right\}.$$

The function h is said to be subdifferentially regular at x if $\partial h(x) = \hat{\partial} h(x)$ and $\partial^\infty h(x) = \hat{\partial} h(x)^\infty$. Subdifferential regularity is important for many reasons, but, in particular, it allows the development of a rich subdifferential calculus.

Let $f : \mathbb{C} \to \overline{\mathbb{R}}$. Define $\Theta : \mathbb{R}^2 \to \mathbb{C}$ by $\Theta(x_1, x_2) = x_1 + \mathrm{i}x_2$ and $\tilde{f} : \mathbb{R}^2 \to \overline{\mathbb{R}}$ by $\tilde{f} = f \circ \Theta$. We say that f is *differentiable in the real sense* if $\tilde{f}$ is differentiable, and f is *twice differentiable in the real sense* if $\tilde{f}$ is twice differentiable. The chain rule gives $f'(\zeta) = \Theta \nabla \tilde{f}(\Theta^{-1}\zeta)$ and $f''(\zeta)\delta = \Theta \nabla^2 \tilde{f}(\Theta^{-1}\zeta)\Theta^{-1}\delta$. Differentiability in the real sense is the only notion of differentiability used in this paper, so we will simply say f is differentiable to mean f is differentiable in the real sense. Let $\langle \cdot, \cdot \rangle$ denote the standard real inner product on $\mathbb{C} : \langle u, v \rangle = \mathrm{Re}[\bar{u}v]$. Then the directional derivative of f in the direction δ is given by $f'(\zeta; \delta) = \langle f'(\zeta), \delta \rangle$, and the second derivative is given by $f''(\zeta; \omega, \delta) = \langle \omega, f''(\zeta)\delta \rangle$. We say that f is quadratic if $f''(\zeta)$ is constant in ζ. For example, the function $r_2(\zeta) = \frac{1}{2}|\zeta|^2$ studied in Section 7 is quadratic with $f'(\zeta) = \zeta$ and $f''(\zeta) = I$.

Finally, we define the *elementary polynomials* $e_{(l,\lambda_0)} \in \mathcal{P}^n$ by $e_{(l,\lambda_0)}(\lambda) = (\lambda - \lambda_0)^l,\ l = 0, \ldots, n$, and recall that, for each fixed value of $\lambda_0 \in \mathbb{C}$, these polynomials form a basis for the linear space $\mathcal{P}^n$.

2. Polynomial root functions

A *max polynomial root functions* is any function of the form

$$\mathbf{f}(p) = \max\{f(\lambda) \mid \lambda \in \mathcal{R}(p)\}, \tag{4}$$

where it is assumed that $f : \mathbb{C} \to \overline{\mathbb{R}}$ is proper, convex, and lsc. We say that $f : \mathbb{C} \to \overline{\mathbb{R}}$ *generates* the max polynomial root function $\mathbf{f} : \mathcal{P}^n \to \overline{\mathbb{R}}$. Two examples of max polynomial root functions are the abscissa ($f(\zeta) = a(\zeta) = \langle 1, \cdot \rangle$) and radius ($f(\zeta) = r(\zeta) = |\zeta|$) mappings on $\mathcal{P}^n$. In [9], the Gauss–Lucas Theorem is used to compute the tangent cone to the epigraph of $\mathbf{f}$ at the polynomial $(\lambda - \lambda_0)^n$. We extend this result by computing the tangent and normal cone to the epigraph at arbitrary monic polynomials and establish the subdifferential regularity of $\mathbf{f}$ under general conditions on the generating function f. In this section, we review the fundamentals required for our development.

Let $\mathcal{P}^n$ denote the linear space of polynomials over $\mathbb{C}$ of degree less than or equal to n, $\mathcal{P}^{n,k} \subset \mathcal{P}^n$ be the subspace of polynomials of degree at most k, and $\mathcal{M}^{n,k} \subset \mathcal{P}^n$ be the subset of polynomials of degree k, for $k = 0, 1, 2, \ldots, n$. Note that $\mathcal{M}^{n,k} \subset \mathcal{P}^{n,k}$, and by $\mathcal{P}^{n,k} \setminus \mathcal{M}^{n,k}$ we mean the relative complement with respect to $\mathcal{P}^{n,k}$. With this notation, $\mathcal{P}^{n,n} = \mathcal{P}^n$. In the relative topology, the set $\mathcal{M}^{n,k}$ is a relatively open dense subset of $\mathcal{P}^{n,k}$ for each k, and $\mathcal{P}^{n,0} \subset \mathcal{P}^{n,1} \subset \cdots \subset \mathcal{P}^{n,n}$. For each $k = 0, 1, 2, \ldots, n$, let $\mathcal{M}_1^{n,k}$ be the set of monic degree k polynomials. The collection $\{\mathcal{M}^{n,0}, \mathcal{M}^{n,1}, \ldots, \mathcal{M}^{n,n}\}$ forms a partition of $\mathcal{P}^n$. When $k = n$, we simplify the notation by setting $\mathcal{M}^n = \mathcal{M}^{n,n}$ and $\mathcal{M}_1^n = \mathcal{M}_1^{n,n}$.

A *weak polynomial root function* (*weak prf*) $\mathbf{h} : \mathcal{P}^n \to \overline{\mathbb{R}}$ is a proper function that is invariant under multiplication by nonzero complex numbers, that is, $\mathbf{h}(p) = \mathbf{h}(\kappa p)$ for all $p \in \mathcal{P}^n$ and for all $\kappa \in \mathbb{C} \setminus \{0\}$. We say that $\mathbf{h}$ is *factor-dominating* at $p \in \mathrm{dom}(\mathbf{h})$ if $\mathbf{h}(q) \le \mathbf{h}(p)$ whenever q divides p and $\deg(q) \ge 1$. If $\mathbf{h}$ is factor-dominating at every $p \in \mathcal{P}^n \setminus \mathcal{M}^{n,0}$, we say $\mathbf{h}$ is factor-dominating. Max polynomial root functions are examples of factor-dominating weak polynomial root functions.

Example 1 (*Root Product Functions*)**.** For $p \in \mathcal{P}^n \setminus \mathcal{M}^{n,0}$, let $\lambda_1, \lambda_2, \ldots, \lambda_{\deg(p)}$ be the roots of p ordered by decreasing modulus and repeated according to multiplicity. Define $\mathbf{h} : \mathcal{P}^n \to \overline{\mathbb{R}}$ by

$$\mathbf{h}(p) = \max \prod_{i=1}^{\deg(p)} |\lambda_i|.$$

Then $\mathbf{h}$ is a weak prf. Moreover, $\mathbf{h}$ is factor-dominating at every polynomial whose roots lie in the complement of the open unit disk. This function is not factor-dominating in general, e.g. consider $p = (\lambda - 1/2)^2$, then $\mathbf{h}(p) = 1/4 < \mathbf{h}(\lambda - 1/2) = 1/2$.

Although most of the prf's of interest are continuous on $\mathcal{M}^{n,k}$ relative to $\mathcal{P}^{n,k}$ for $k = 1, 2, \ldots, n$, they are not Lipschitz continuous there, nor are they bounded in the neighborhood of any point on the boundary of $\mathcal{M}^{n,k}$ relative to $\mathcal{P}^{n,k}$ for any k. Indeed, this is the case for the polynomial abscissa map $\mathbf{a}$. For example, if $p_\varepsilon(\lambda) = \lambda^n - \varepsilon$, then $\mathbf{a}(p_\varepsilon) = \sqrt[n]{\varepsilon}$ is not

Lipschitz continuous at $\varepsilon = 0$. In addition, given $q \in \mathcal{P}^{n,n-1} \setminus \mathcal{M}^{n,0}$, define $p_\varepsilon(\lambda) = (1 - \varepsilon\lambda)q$ so that $p_\varepsilon \to q$ as $\varepsilon \downarrow 0$. But $\mathbf{a}((1-\varepsilon\lambda)q) = \max\{1/\varepsilon, \mathbf{a}(q)\} \to \infty$ as $\varepsilon \downarrow 0$.

Let $\mathbf{h}_1 : \mathcal{P}^n \to \overline{\mathbb{R}}$ be given by $\mathbf{h}_1 = \mathbf{h} + \delta_{\mathcal{M}_1^n}$, where

$$\delta_{\mathcal{M}_1^n}(p) = \begin{cases} 0 & \text{if } p \in \mathcal{M}_1^n, \\ +\infty & \text{otherwise,} \end{cases}$$

is the convex indicator function of $\mathcal{M}_1^n$. Note that $\mathrm{dom}(\mathbf{h}_1) = \mathrm{dom}(\mathbf{h}) \cap \mathcal{M}_1^n$. We now extend [9, Lemma 1] to weak prf's and arbitrary polynomials in $\mathcal{M}^n \cap \mathrm{dom}(\mathbf{h})$.

Lemma 2.1. *Let* $\mathbf{h}$ *be a weak prf and let* $\mathbf{h}_1$ *be as above. Then*

$$\mathcal{M}^n \cap \mathrm{dom}(\mathbf{h}) = \{\kappa q \mid \kappa \in \mathbb{C} \setminus \{0\},\ q \in \mathcal{M}_1^n \cap \mathrm{dom}(\mathbf{h}_1)\}. \tag{5}$$

Moreover, for $q \in \mathcal{M}_1^n \cap \mathrm{dom}(\mathbf{h}_1)$, $\kappa \in \mathbb{C} \setminus \{0\}$, *and* $p = \kappa q$,

$$T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p)) = \{(\zeta p + \kappa \tilde{v}, \eta) \mid \zeta \in \mathbb{C},\ (\tilde{v}, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q))\} \tag{6}$$

and

$$T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q)) = \left\{ (\kappa^{-1}(v - \omega q), \eta) \mid \begin{array}{l} (v, \eta) \in T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p)) \text{ and } \omega \text{ is the unique element of } \mathbb{C} \\ \text{such that } v - \omega p \in \mathcal{P}^{n-1} \end{array} \right\}. \tag{7}$$

Proof. Observe that, for $r \in \mathcal{M}_1^n$ and $\gamma \in \mathbb{C} \setminus \{0\}$, we have $\mathbf{h}(\gamma r) = \mathbf{h}_1(r)$, and consequently, $(r, \tau) \in \mathrm{epi}(\mathbf{h}_1)$ if and only if $(\gamma r, \tau) \in \mathrm{epi}(\mathbf{h})$ for all $\gamma \in \mathbb{C} \setminus \{0\}$. This proves (5).

It is easily shown that the identities (6) and (7) are equivalent, and so we only prove (6). Given p in $\mathcal{M}^n \cap \mathrm{dom}(\mathbf{h})$, suppose $(v, \eta) \in \mathcal{P}^n \times \mathbb{R}$ is such that $(v, \eta) \in T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p))$. Then there exists $\xi_i \downarrow 0$ such that

$$(p + \xi_i v + o(\xi_i),\ \mathbf{h}(p) + \xi_i \eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}). \tag{8}$$

Let (q, κ) be the unique pair in $\mathcal{M}_1^n \times (\mathbb{C} \setminus \{0\})$ for which $p = \kappa q$ and $\omega \in \mathbb{C}$ be the unique element such that $v - \omega p \in \mathcal{P}^{n-1}$. Set $\tilde{v} = \kappa^{-1}(v - \omega p)$ so that $v = \omega p + \kappa \tilde{v}$. We now show that $(\tilde{v}, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q))$ which implies that $T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p))$ is contained in the set on the right-hand side of (6). To this end let $\hat{o}(\xi)$ be such that $\tilde{o}(\xi_i) = o(\xi_i) - \hat{o}(\xi_i)p \in \mathcal{P}^{n-1}$. Then (8) becomes

$$(\kappa(1 + \omega\xi_i + \hat{o}(\xi_i))q + \kappa\xi_i\tilde{v} + \tilde{o}(\xi_i),\ \mathbf{h}(p) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}),$$

where $\kappa\xi_i\tilde{v} + \tilde{o}(\xi_i) \in \mathcal{P}^{n-1}$. This implies

$$(q + \xi_i\tilde{v}/(1 + \omega\xi_i + \hat{o}(\xi_i)) + \tilde{o}(\xi_i)/\kappa(1 + \omega\xi_i + \hat{o}(\xi_i)),\ \mathbf{h}_1(q) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}_1)$$

for all i sufficiently large. Since $(1 + \xi_i\omega + \hat{o}(\xi_i))^{-1} = 1 + O(\xi_i)$, we have $(\tilde{v}, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q))$.

For the reverse inclusion, suppose $(\tilde{v}, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q))$, let $\zeta \in \mathbb{C}$, and define $v = \zeta p + \kappa\tilde{v}$. By the definition of the tangent cone, there exists $\xi_i \downarrow 0$ such that

$$(q + \xi_i\tilde{v} + o(\xi_i),\ \mathbf{h}_1(q) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}_1).$$

Substituting $\tilde{v} = \kappa^{-1}(v - \zeta p)$ gives

$$(q + \xi_i\kappa^{-1}(v - \zeta p) + o(\xi_i),\ \mathbf{h}_1(q) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}_1),$$

that is,

$$((1 - \xi_i\zeta)q + \kappa^{-1}\xi_i v + o(\xi_i),\ \mathbf{h}_1(q) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}_1).$$

Multiplying by κ gives

$$((1 - \xi_i\zeta)p + \xi_i v + o(\xi_i),\ \mathbf{h}(p) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}).$$

Since $(1 - \xi_i\zeta) \neq 0$ for i sufficiently large, this implies

$$(p + \xi_i v/(1 - \xi_i\zeta) + o(\xi_i),\ \mathbf{h}(p) + \xi_i\eta + o(\xi_i)) \in \mathrm{epi}(\mathbf{h}).$$

Thus, $(v, \eta) \in T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p))$, which concludes the proof of (6). □

In light of the equivalences (6) and (7), we need only compute a representation for the tangent cone $T_{\mathrm{epi}(\mathbf{h}_1)}(q, \mathbf{h}_1(q))$ in order to obtain one for $T_{\mathrm{epi}(\mathbf{h})}(p, \mathbf{h}(p))$. This simplifies the derivations since it allows us to restrict the analysis to the affine manifold $\mathcal{M}_1^n$.

We now provide a formal definition for polynomial root functions. Let $\preceq$ denote the *lexicographical order* on $\mathbb{C}$ where for $z_s = x_s + \mathrm{i}y_s,\ x_s, y_s \in \mathbb{R},\ s = 1, 2$, we have $z_1 \preceq z_2$ if and only if either $x_1 < x_2$ or ($x_1 = x_2$ and $y_1 \le y_2$). For a polynomial

$p \in \mathcal{P}^n \setminus \mathcal{M}^{n,0}$ of degree k, we label its roots $\lambda_1, \lambda_2, \dots, \lambda_k$ according to the lexicographic ordering and repeated according to multiplicity. Next define the family of maps $\zeta_k : \mathcal{P}^k \to \mathbb{C}^k$ $k = 0, 1, \dots, n$ by

$$\zeta_k(p) = (\lambda_1, \lambda_2, \dots, \lambda_k)^T$$

when $k = \deg(p) \geq 1$ and $\zeta_0(p) = 0$ when $\deg(p) = 0$. We write $\zeta(p)$ to mean $\zeta_{\deg(p)}(p)$ and suppress the subscript $\deg(p)$. Consider a family of functions $h_k : \mathbb{C}^k \to \overline{\mathbb{R}}$, $k = 0, 1, 2, \dots, n$, such that each h_k is invariant under permutations of its arguments and $h_0 : \{0\} \to \overline{\mathbb{R}}$ is identically $+\infty$. We define the associated family $\mathbf{h}_k : \mathcal{M}^{n,k} \to \overline{\mathbb{R}}$ by $\mathbf{h}_k = h_k \circ \zeta_k$, and $\mathbf{h} : \mathcal{P}^n \to \overline{\mathbb{R}}$ by $\mathbf{h}(p) = \mathbf{h}_k(p)$, where $k = \deg(p)$. More simply, we write $\mathbf{h} = h \circ \zeta$, where we suppress the subscripts and the choice of the family $\{h_0, h_1, \dots, h_n\}$. We call $\mathbf{h}$ a *polynomial root function* or prf. A polynomial root function is always a weak polynomial root function. All of the polynomial root functions $\mathbf{h} : \mathcal{P}^n \to \overline{\mathbb{R}}$ we consider have the property that $\mathbf{h}|_{\mathcal{M}^{n,k} \cap \mathrm{dom}(\mathbf{h})}$ is continuous for $k = 1, 2, \dots, n$.

2.1. Factorization spaces

Factorization spaces [9] are used to extend facts about polynomials of the form $(\lambda - \lambda_0)^n$ to general polynomials. Let $(n_1, n_2, \dots, n_m)$ be a partition of n and let $p_j \in \mathcal{M}_1^{n_j}$, $i = 1, \dots, m$ be relatively prime as elements of $\mathcal{P}^n$. Define the *factorization space* for the polynomials

$$\pi = (p_1, p_2, \dots, p_m) \tag{9}$$

to be the product space

$$\mathcal{S}_\pi = \mathbb{C} \times \mathcal{P}^{n_1 - 1} \times \mathcal{P}^{n_2 - 1} \times \cdots \times \mathcal{P}^{n_m - 1}. \tag{10}$$

The component indexing for elements of $\mathcal{S}_\pi$ starts with zero so that the jth component is an element of $\mathcal{P}^{n_j - 1}$. If $\pi = (e_{(n_1,\lambda_1)}, \dots, e_{(n_m,\lambda_m)})$ is *the* prime factorization for $p \in \mathcal{M}_1^n$, where

$$p = \prod_{j=1}^{m} e_{(n_j,\lambda_j)}, \tag{11}$$

with $\lambda_1, \lambda_2, \dots, \lambda_m$ the distinct roots of p ordered lexicographically, then we write

$$\mathcal{S}_p = \mathbb{C} \times \mathcal{P}^{n_1 - 1} \times \mathcal{P}^{n_2 - 1} \times \cdots \times \mathcal{P}^{n_m - 1}. \tag{12}$$

The spaces $\mathcal{P}^n$ and $\mathcal{S}_\pi$ are related through the mapping $F_\pi : \mathcal{S}_\pi \to \mathcal{P}^n$ given by

$$F_\pi(q_0, q_1, q_2, \dots, q_m) = (1 + q_0) \prod_{j=1}^{m} (p_j + q_j). \tag{13}$$

Note that $F_\pi(0) = \prod_{j=1}^{m} p_j$, and since the polynomials in (9) are relatively prime, [2, Lemma 1.4] tells us that there exist neighborhoods U of 0 in $\mathcal{S}_\pi$ and V of $\prod_{j=1}^{m} p_j$ in $\mathcal{P}^n$ such that $F_\pi|_U : U \to V$ is a diffeomorphism. Thus, $\nabla F_\pi(0) : \mathcal{S}_\pi \to \mathcal{P}^n$, given by

$$\nabla F_\pi(0)(\omega_0, w_1, w_2, \dots, w_m) = \omega_0 \prod_{j=1}^{m} p_j + \sum_{j=1}^{m} r_j w_j,$$

is an isomorphism, where $r_j = \prod_{s \neq j} p_s$ for $j = 1, 2, \dots, m$.

Let

$$\mathbf{h}_1 = \mathbf{h} + \delta_{\mathcal{M}_1^n} \tag{14}$$

and define $\mathbf{h}_{[1,n_j]} : \mathcal{P}^{n_j} \to \overline{\mathbb{R}}$ by

$$\mathbf{h}_{[1,n_j]}(q) = \begin{cases} \mathbf{h}(q) & \text{if } q \in \mathcal{M}_1^{n_j} \\ +\infty & \text{otherwise.} \end{cases} \tag{15}$$

Following the approach taken in [2] for the abscissa mapping, we show that if $\mathbf{h}$ is a factor-dominating prf, then the tangent cone

$$T_{\mathrm{epi}(\mathbf{h}_1)}\left(\prod_{j=1}^{m} p_j, \mathbf{h}_1\left(\prod_{j=1}^{m} p_j\right)\right)$$

can be decomposed into a kind of *product* of the tangent cones $T_{\mathrm{epi}(\mathbf{h}_{[1,n_j]})}(p_j, \mathbf{h}_1(p_j))$.

Theorem 2.2. *Let* $\mathbf{h}$ *be a factor-dominating prf, let* $\mathbf{h}_1$ *and* $\mathbf{h}_{[1,n_j]}$ *be as in* (14) *and* (15)*, let* $\pi = (p_1, p_2, \dots, p_m)$ *be as in* (9) *with* $\prod_{j=1}^m p_j \in \mathrm{dom}(\mathbf{h}) \cap \mathcal{M}_1^n$*, and let* $\mathcal{S}_\pi$ *be as in* (10)*. If*

$$(v, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}\left(\prod_{j=1}^m p_j, \mathbf{h}\left(\prod_{j=1}^m p_j\right)\right),$$

then there exists $(0, w_1, w_2, \dots, w_m) \in \mathcal{S}_\pi$ *such that* $v = \sum_{j=1}^m r_j w_j$*, with*

$$(w_j, \eta) \in T_{\mathrm{epi}(\mathbf{h}_{[1,n_j]})}\left(p_j, \mathbf{h}\left(\prod_{j=1}^m p_j\right)\right),$$

for $j = 1, 2, \dots, m$.

Proof. Set $p = \prod_{j=1}^m p_j$. Let $(v, \eta) \in T_{\mathrm{epi}(\mathbf{h}_1)}(p, \mathbf{h}_1(p))$ with $v \neq 0$. Then there exist $\xi_k \downarrow 0$ and sequences $\{o_{1k}\} \subset \mathcal{P}^{n-1}$ and $\{o_{2k}\} \subset \mathbb{C}$ such that $o_{1k}/\xi_k \to 0$, $o_{2k}/\xi_k \to 0$, and

$$(p + \xi_k v + o_{1k},\ \mathbf{h}(p) + \xi_k \eta + o_{2k}) \in \mathrm{epi}(\mathbf{h}_1), \quad k = 1, 2, \dots.$$

Set $q_k = p + \xi_k v + o_{1k}$, $k = 1, 2 \dots$. Since $\mathrm{dom}(\mathbf{h}) \subset \mathcal{M}^n$ and F_π is a local diffeomorphism at the origin, there is a constant $K > \|\nabla(F_\pi^{-1}(0))\|$ and a sequence $\{(0, u_{1k}, u_{2k}, \dots, u_{mk})\} \subset \mathcal{S}_\pi$ such that $q_k = p + \xi_k v + o_{1k} = \prod_{k=1}^m (p_j + u_{jk})$, with $\|(0, u_{1k}, u_{2k}, \dots, u_{mk})\| \le K\|q_k - p\| \le K(\xi_k\|v\| + \|o_{1k}\|)$. Hence, by passing to a subsequence if necessary, we can assume with no loss in generality that there exist $(0, w_1, w_2, \dots, w_m) \in \mathcal{S}_\pi$ such that

$$\xi_k^{-1}(0, u_{1k}, u_{2k}, \dots, u_{mk}) \to (0, w_1, w_2, \dots, w_m),$$

or equivalently,

$$(0, u_{1k}, u_{2k}, \dots, u_{mk}) = \xi_k(0, w_1, w_2, \dots, w_m) + o_{3k},$$

where $o_{3k}/\xi_k \to 0$ and $o_{3k} = (0, o_{3k1}, \dots, o_{3km})$. But, since $\mathbf{h}$ is factor-dominating, so is $\mathbf{h}_1$, giving

$$(p_j + u_{jk}, \mathbf{h}(p) + \xi\eta + o_{2k}) \in \mathrm{epi}(\mathbf{h}_{[1,n_j]}) \quad \text{for all } j = 1, \dots, m \text{ and } k = 1, 2, \dots.$$

Consequently, for $j = 1, 2, \dots, m$, $(w_j, \eta) \in T_{\mathrm{epi}(\mathbf{h}_{[1,n_j]})}(p_j, \mathbf{h}(p))$. In addition,

$$p + \xi_k v + o_{1k} = \prod_{k=1}^m (p_j + \xi_k w_j + o_{3kj}) = p + \xi_k \sum_{k=1}^m w_j r_j + o_k,$$

where $o_k/\xi_k \to 0$. Therefore, $v = \sum_{k=1}^m w_j r_j$ which proves the result. □

3. Subderivative and tangent cone

We now focus our attention on the max polynomial root functions $\mathbf{f}$ defined in (4). We begin with the formula for the subderivative $\mathbf{df}(e_{(n,\lambda_0)})$ given in [9, Theorem 6]. This result, as well as many of those that follow, makes use of one or the other of the following two assumptions:

(A) f is twice continuously differentiable at λ and $f''(\lambda; \cdot, \cdot)$ is positive definite or f is quadratic,
(B) $\mathrm{rspan}\,(\partial f(\lambda)) = \mathbb{C}$,

where $\mathrm{rspan}\,(\partial f(\lambda)) = \{\tau\zeta \mid \tau \in \mathbb{R},\ \zeta \in \partial f(\lambda)\}$ is the real linear span of the set $\partial f(\lambda)$.

Theorem 3.1 *([9, Theorem 6]). Let* $\lambda_0 \in \mathrm{dom}(\partial f)$ *be such that* $\partial f(\lambda_0) \neq \{0\}$*, and let* $v \in \mathcal{P}^n$ *be such that* $v = \sum_{k=0}^n \omega_k e_{(k,\lambda_0)}$*. If any one of the conditions*

$$0 = \left\langle g, \sqrt{-\omega_2}\right\rangle \quad \text{for all } g \in \partial f(\lambda_0), \tag{16}$$

$$\omega_k = 0 \quad \text{for all } k = 3, \dots, n, \tag{17}$$

is not satisfied, then $\mathbf{df}(e_{(n,\lambda_0)})(v) = +\infty$*; otherwise,*

$$\mathbf{df}(e_{(n,\lambda_0)})(v) \ge f'(\lambda_0; -\omega_1)/n, \tag{18}$$

with equality holding if (B) *is satisfied with* $\lambda = \lambda_0$*. If* f *satisfies* (A) *at* $\lambda = \lambda_0$*, then*

$$\mathbf{df}(e_{(n,\lambda_0)})(v) = (f'(\lambda_0; -\omega_1) + f''(\lambda_0; \sqrt{-\omega_2}, \sqrt{-\omega_2}))/n, \tag{19}$$

whenever (16) *and* (17) *hold. Moreover, if either* (A) *or* (B) *is satisfied, then the subderivative* $\mathbf{df}(e_{(n,\lambda_0)})$ *is proper, sublinear, and lsc.*

Remark 1. The requirement that $\partial f(\lambda_0) \neq \{0\}$ is used to obtain the conditions (17).

Remark 2. Under the convention that

$$f''(\lambda_0; \cdot, \cdot) = 0 \quad \text{whenever } f''(\lambda_0) \text{ does not exist,} \tag{20}$$

equality in (18) under (B) is equivalent to (19).

Proof. Conditions (18) and (19) follow from [9, Theorem 6]. Since $\lambda_0 \in \mathrm{dom}(f)$, we have that $(0, 1) \in T_{\mathrm{epi}(\mathbf{f}_1)}(e_{(n,\lambda_0)}, f(\lambda_0))$. Therefore, $\mathbf{df}_1(e_{(n,\lambda_0)})(0) = 0$, so the subderivative is proper. The subderivative is always positively homogeneous, so we need only show it is subadditive. Let

$$v^k = \sum_{s=1}^{n} \omega_s^k e_{(n-s,\lambda_0)}, \quad k = 1, 2, \text{ so that } v^1 + v^2 = \sum_{s=1}^{n} (\omega_s^1 + \omega_s^2) e_{(n-s,\lambda_0)}.$$

Clearly $\mathbf{df}_1(e_{(n,\lambda_0)})(v^1 + v^2) \leq \mathbf{df}_1(e_{(n,\lambda_0)})(v^1) + \mathbf{df}_1(e_{(n,\lambda_0)})(v^2)$ if either v^1 or v^2 violates either (16) or (17) since then $\mathbf{df}_1(e_{(n,\lambda_0)})(v^1) + \mathbf{df}_1(e_{(n,\lambda_0)})(v^2) = +\infty$. Therefore, we assume that both v^1 and v^2 satisfy (16) and (17). It is easily verified that given $a, b \in \mathbb{C}$, $\langle a, \sqrt{-b} \rangle = 0$ if and only if either $a = 0$ or $b = ta^2$ for some $t \geq 0$. Hence, given $g \in \partial f(\lambda_0)$, condition (16) implies that either $g = 0$ or there exists $t_1, t_2 \geq 0$ such that $\omega_2^k = t_k g^2$, $k = 1, 2$. Therefore, $\omega_2^1 + \omega_2^2 = (t_1 + t_2) g^2$ and $\langle g, \sqrt{-(\omega_2^1 + \omega_2^2)} \rangle = \sqrt{t_1 + t_2}\,\mathrm{Re}(\mathrm{i}|g|^2) = 0$, that is, $v^1 + v^2$ also satisfies (16). Hence,

$$n\,\mathbf{df}_1(e_{(n,\lambda_0)})(v^1 + v^2) = f'(\lambda_0; -(\omega_1^1 + \omega_1^2)) + f''\left(\lambda_0; \sqrt{-(\omega_2^1 + \omega_2^2)}, \sqrt{-(\omega_2^1 + \omega_2^2)}\right),$$

where the first term on the right-hand side is sublinear since f is convex and the second term is sublinear from [9, Lemma 5] (here we use the convention (20)). $\square$

The final statement of Theorem 3.1 concerning the sublinearity of the subderivative $\mathbf{df}_1(e_{(n,\lambda_0)})$ does not appear in [9, Theorem 6]. This addition is the cornerstone to our derivation of the subderivative $\mathbf{df}_1(p)$ for general monic polynomials p. The sublinearity of $\mathbf{df}_1(e_{(n,\lambda_0)})$ in conjunction with (2) implies that a representation for the regular subdifferential can be obtained by representing the right-hand side of (19) as a support function. This is the first step in the derivation of the regular subdifferential in Section 5. In the remainder of this section we extend Theorem 3.1 to general polynomials in $\mathcal{M}_1^n$.

Suppose $p \in \mathcal{M}_1^n$ has prime factorization (11). In [2, Theorem 1.6], the factorization space $\mathcal{S}_p$ (12) is used to decompose the tangent cone at $(p, \mathbf{f}(p))$ into a kind of product of the tangent cones of the form $T_{\mathrm{epi}(\mathbf{f}_{[1,n_j]})}(e_{(n_j,\lambda_j)}, \mathbf{f}(p))$, where $\mathbf{f}_{[1,n_j]} : \mathcal{P}^{n_j} \to \overline{\mathbb{R}}$ is given as in (15) by

$$\mathbf{f}_{[1,n_j]}(q) = \begin{cases} \mathbf{f}(q) & \text{if } q \in \mathcal{M}_1^{n_j} \\ +\infty & \text{otherwise.} \end{cases} \tag{21}$$

Theorem 2.2 gives necessary conditions for $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$ in terms of the prime factorization (11). In the following result we use assumptions (A) and (B) to show that these conditions are also sufficient.

Theorem 3.2. *Let f be proper, lsc and convex. Let $p \in \mathrm{dom}(\mathbf{f}) \cap \mathcal{M}_1^n$ be as in (11) and define $\mathcal{I}(p) = \{j \mid \mathbf{f}(p) = f(\lambda_j),\ j = 1, \ldots, m\}$. If $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$, then there exists a point $(\omega_0, w_1, \ldots, w_m) \in \mathcal{S}_p$ satisfying*

$$\omega_0 = 0 \tag{22}$$

$$v = \nabla F_p(0)(\omega_0, w_1, \ldots, w_m) = \sum_{j=1}^{m} r_j w_j, \quad \text{where } r_j = \prod_{k \neq j} e_{(n_k,\lambda_k)} \tag{23}$$

$$(w_j, \eta) \in T_{\mathrm{epi}(\mathbf{f}_{[1,n_j]})}(e_{(n_j,\lambda_j)}, f(\lambda_j)) \quad \text{for } j \in \mathcal{I}(p), \tag{24}$$

$$(w_j, \eta) \in T_{\mathrm{epi}(\mathbf{f}_{[1,n_j]})}(e_{(n_j,\lambda_j)}, \mathbf{f}(p)) \quad \text{for } j \notin \mathcal{I}(p). \tag{25}$$

These conditions are sufficient for (v, η) to be an element of $T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$ if $\partial f(\lambda_j) \neq \{0\}$ and f satisfies either (A) *or* (B) *at $\lambda = \lambda_j$ for every $\lambda_j \in \mathcal{I}(p)$. In this case,* epi$(\mathbf{f}_1)$ *is geometrically derivable at p.*

Proof. That (22)–(25) are necessary for $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$ follows from Theorem 2.2. So we need only establish the sufficiency of (22)–(25). Let $(v, \eta) \in \mathcal{P}^{n-1} \times \mathbb{R}$ be such that $v = \sum_{j=1}^{m} r_j w_j$ satisfies (24)–(25). As in the proof of [9, Theorem 6], we use a carefully chosen trajectory of polynomials, $\gamma(\xi) = (p_\xi, \mathbf{f}(p_\xi))$, satisfying $\gamma'(0) = (v, \eta)$ showing that $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$ and that epi $(\mathbf{f}_1)$ is geometrically derivable. The trajectory is built up from factors of the form

$$\begin{aligned} q(\lambda; \lambda_0, \xi, \varphi, k, l, w) &= \left(\lambda - (\lambda_0 - (\xi/k)(\omega_1 - \varphi/(2l)) + \sqrt{-\omega_2 \xi / l})\right)^l \\ &\quad \times \left(\lambda - (\lambda_0 - (\xi/k)(\omega_1 - \varphi/(2l)) - \sqrt{-\omega_2 \xi / l})\right)^l \\ &= (\lambda - \lambda_0)^{2l} + (2l\xi/k)(\omega_1 - \varphi/(2l))(\lambda - \lambda_0)^{2l-1} + \omega_2 \xi (\lambda - \lambda_0)^{2l-2} + o(\xi), \end{aligned}$$

where $\lambda_0 \in \mathbb{C}, \xi > 0, \varphi \in \mathbb{C}, k \in \{1,\dots,n\}, l \in \{1,\dots,k/2\}$, and $w = \sum_{s=1}^{k} \omega_s e_{(k-s,\lambda_0)} \in \mathcal{P}^{k-1}$, one for each of the polynomials p_j, $j = 1,\dots,m$.

Let $j \in \mathcal{I}(p)$. By [9, Theorem 7] (or, equivalently, by combining (1) and Theorem 3.1),

$$w_j(\lambda) = \omega_{j1}(\lambda-\lambda_j)^{n_j-1} + \omega_{j2}(\lambda-\lambda_j)^{n_j-2},$$

where $\omega_{j1}, \omega_{j2} \in \mathbb{C}$ (with $\omega_{j2} = 0$ if $n_j = 1$) satisfy

$$\eta \geq (1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] \quad \text{and} \tag{26}$$

$$0 = \left\langle f'(\lambda_j), \sqrt{-\omega_{j2}} \right\rangle, \tag{27}$$

with $f''(\lambda_0; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}}) = 0$ if assumption (B) holds.

First consider the case where λ_j satisfies assumption (A). There are two sub-cases to consider: n_j is even and n_j is odd.

If n_j is even set $l_j = n_j/2$. For $\xi > 0$, define $p_{\xi,j}(\lambda) = q(\lambda; \lambda_j, \xi, 0, n_j, l_j, w_j)$. The roots of $p_{\xi,j}$ are $\lambda_j - (\xi/n_j)\omega_{j1} \pm \sqrt{-\omega_{j2}\xi/l_j}$, and so

$$\mathbf{f}_{[1,n_j]}(p_{\xi,j}) = \max\left\{ f(\lambda_j - (\omega_{j1}/n_j)\xi + \sqrt{-\omega_{j2}\xi/l_j}), f(\lambda_j - (\omega_{j1}/n_j)\xi - \sqrt{-\omega_{j2}\xi/l_j}) \right\}.$$

Using the second-order Taylor expansion of f about λ_j and (27), these roots yield

$$\mathbf{f}_{[1,n_j]}(p_{\xi,j}) = f(\lambda_j) + (\xi/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] + o_j(\xi). \tag{28}$$

If n_j is odd let l_j be such that $n_j = 2l_j + 1$, and define

$$\varphi_j = -\frac{f''\left(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}}\right)}{\overline{f'(\lambda_j)}}.$$

The scalars φ_j are well defined since we have assumed that $f'(\lambda_j) \neq 0$ for every $j \in \mathcal{I}(p)$. For $\xi > 0$ set $p_{\xi,j}(\lambda) = q(\lambda; \lambda_j, \xi, \varphi_j, n_j, l_j, w_j)$. The roots of $p_{\xi,j}$ are

$$\lambda_j - (\xi/n_j)(\omega_{j1} + \varphi_j) \quad \text{and} \quad \lambda_j - (\xi/n_j)\left(\omega_{j1} - \varphi_j/(2l_j)\right) \pm \sqrt{-\omega_{j2}\xi/l_j},$$

and so

$$\begin{aligned}\mathbf{f}_{[1,n_j]}(p_{\xi,j}) = \max\Big\{ & f(\lambda_j - (\xi/n_j)(\omega_{j1}+\varphi_j)), f(\lambda_j - (\xi/n_j)\left(\omega_{j1} - \varphi_j/(2l_j)\right) + \sqrt{-\omega_{j2}\xi/l_j},), \\ & f(\lambda_j - (\xi/n_j)\left(\omega_{j1} - \varphi_j/(2l_j)\right) - \sqrt{-\omega_{j2}\xi/l_j},) \Big\}.\end{aligned}$$

Again, by taking the second-order Taylor expansion of f at λ_j and using (27) with the definition of φ_j, these roots yield the equivalence (28).

Next, suppose $j \in \mathcal{I}(p)$ is such that λ_j satisfies assumption (B) instead of (A). In this case define $p_{\xi,j}(\lambda) = (\lambda - (\lambda_j - \xi\omega_{1j}/n_j))^{n_j}$. Using (27) and convention (20), again gives the equivalence (28).

Therefore, (28) holds for all $j \in \mathcal{I}(p)$. Consequently, $p_{\xi,j} \in \mathrm{dom}(\mathbf{f})$ for all $j \in \mathcal{I}(p)$ and ξ sufficiently small, with

$$\mathbf{f}_{[1,n_j]}(p_{\xi,j}) = f(\lambda_j) + (\xi/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] + o(\xi) \leq f(\lambda_j) + \xi\eta + o_j(\xi),$$

where the inequality follows from (26).

If $j \notin \mathcal{I}(p)$ define $p_{\xi,j}(\lambda) = (\lambda - \lambda_j)^{n_j} + \xi w_j$ and set $p_\xi = \prod_{j=1}^{m} p_{\xi,j}$. By the continuity of $\mathbf{f}$ on $\mathcal{M}_1^{n,n_j}$, we have that $p_{\xi,j} \in \mathrm{dom}(\mathbf{f}_{[1,n_j]})$, which implies that $p_\xi \in \mathrm{dom}(\mathbf{f})$ and for each $j_0 = 1, 2, \dots, m$,

$$\mathbf{f}_{[1,n_{j_0}]}(p_{\xi,j_0}) \leq \max_{j\in\mathcal{I}(p)}\{\mathbf{f}_{[1,n_j]}(p_{\xi,j})\} \leq \max_{j\in\mathcal{I}(p)}\{f(\lambda_j) + \xi\eta + o_j(\xi)\}$$

for ξ sufficiently small. Set $\beta_\xi = \max_{j\in\mathcal{I}(p)}\{f(\lambda_j) + \eta\xi + o_j(\xi)\}$. Then for small ξ, $(p_\xi, \beta_\xi) \in \mathrm{epi}(\mathbf{f}_1)$. That $(\beta_\xi - \mathbf{f}(p))/\xi \to \eta$ as $\xi \downarrow 0$ follows immediately from the definition of the sequence β_ξ. Also,

$$\begin{aligned}(p_\xi - p)/\xi &= (F_p(0) + \xi\nabla F(0)(0, w_1, w_2, \dots, w_m) + o(\xi) - F_p(0))/\xi \\ &= \nabla F_p(0, w_1, \dots, w_m) + o(\xi)/\xi \\ &\to \nabla F_p(0, w_1, \dots, w_m) = v \quad \text{as } \xi \downarrow 0.\end{aligned}$$

Therefore, $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f}_1)}(p, \mathbf{f}(p))$. □

We now describe the subderivative of $\mathbf{f}$ at $p \in \mathcal{M}^n$.

Theorem 3.3. *Suppose f is proper, convex and lsc. Let $p \in \mathrm{dom}(\mathbf{f}) \cap \mathcal{M}_1^n$ be as in (11) with $\partial f(\lambda_j) \neq \{0\}$ for all $j \in \mathcal{I}(p)$. If $v = \nabla F_p(0)(\omega_0, w_1, \ldots, w_m)$ with $w_j = \sum_{s=1}^{n_j} \omega_{js} e_{(n_j-s,\lambda_j)}(\lambda)$ for $j = 1, 2, \ldots, m$ satisfies*

$$0 = \left\langle g, \sqrt{-\omega_{j2}} \right\rangle \quad \text{for all } g \in \partial f(\lambda_j),\ j \in \mathcal{I}(p), \tag{29}$$

$$\text{and} \quad 0 = \omega_{jk}, \quad k = 3, \ldots, n_j \text{ for all } j \in \mathcal{I}(p), \tag{30}$$

then

$$\mathrm{d}\mathbf{f}(p)(v) \geq \max_{j \in \mathcal{I}(p)} \{f'(\lambda_j; -\omega_{j1})/n_j\}; \tag{31}$$

otherwise, $\mathrm{d}\mathbf{f}(p)(v) = +\infty$*. If, in addition, f satisfies either* (A) *or* (B) *at $\lambda = \lambda_j$ for all $j \in \mathcal{I}(p)$, then*

$$\mathrm{d}\mathbf{f}(p)(v) = \max_{j \in \mathcal{I}(p)} \{[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})]/n_j\}, \tag{32}$$

whenever $(\omega_0, w_1, \ldots, w_m)$ satisfy (29) and (30) for all $j \in \mathcal{I}(p)$, where we use the convention (20) if (B) *holds. In this case, $\mathrm{d}\mathbf{f}(p)$ is proper, lsc, and sublinear.*

Proof. Inequality (31) follows directly from Theorems 3.1 and 3.2, so we only discuss equality (32). Note that $(v, \mathrm{d}\mathbf{f}(p)(v)) = (v, \eta)$ for some $(v, \eta) \in T_{\mathrm{epi}(\mathbf{f})}(p, \mathbf{f}(p))$. By (26),

$$\mathrm{d}\mathbf{f}(p)(v) \geq \max_{j \in \mathcal{I}(p)} \{(1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})]\}.$$

Let $p_{\xi,j}, p_\xi = \prod_{j=1}^m p_{\xi,j}$ be as in the proof of Theorem 3.2. We have $p_\xi = p + \xi v + o(\xi)$. Provided $\lim_{\xi \downarrow 0}(\mathbf{f}(p_\xi) - \mathbf{f}(p))/\xi$ exists, we have $\mathrm{d}\mathbf{f}(p)(v) = \liminf_{\xi \downarrow 0, q \to v}(\mathbf{f}(p + \xi q) - \mathbf{f}(p))/\xi \leq \lim_{\xi \downarrow 0}(\mathbf{f}(p_\xi) - \mathbf{f}(p))/\xi$. By (28),

$$\mathbf{f}(p_\xi) = \max_{j=1,2,\ldots,m} \{f(\lambda_j) + \xi(1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] + o(\xi)\}.$$

Since $\xi > 0$, $(\mathbf{f}(p_\xi) - \mathbf{f}(p))/\xi$ equals

$$\begin{aligned}&\max\{\{(1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] + o(\xi)/\xi\}_{j \in \mathcal{I}(p)},\\ &\quad \{(f(\lambda_j) - \mathbf{f}(p))/\xi + (1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})] + o(\xi)/\xi\}_{j \notin \mathcal{I}(p)}\}.\end{aligned}$$

Furthermore, $\{(1/n_j)f'(\lambda_j; -\omega_{j1})\}_{j=1,2,\ldots,m}$ and $\{(1/n_j)f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})\}_{j=1,2,\ldots,m}$ are bounded, and $(f(\lambda_j) - \mathbf{f}(p))/\xi$ is strictly negative and bounded away from zero for all $j \notin \mathcal{I}(p)$. So for small ξ,

$$(\mathbf{f}(p_\xi) - \mathbf{f}(p))/\xi = \max_{j \in \mathcal{I}(p)} \{[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})]/n_j + o(\xi)/\xi\}.$$

Therefore, $\lim_{\xi \downarrow 0}(\mathbf{f}(p_\xi) - \mathbf{f}(p))/\xi = \max_{j \in \mathcal{I}(p)}\{(1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})]\}$, which implies that $\mathrm{d}\mathbf{f}(p)(v) \leq \max_{j \in \mathcal{I}(p)}\{(1/n_j)[f'(\lambda_j; -\omega_{j1}) + f''(\lambda_j; \sqrt{-\omega_{j2}}, \sqrt{-\omega_{j2}})]\}$. By Theorem 3.1, each term in the maximum in the above display is proper, sublinear and lsc. Therefore $\mathrm{d}\mathbf{f}(p)(\cdot)$ is proper, sublinear and lsc. □

When $\mathrm{d}\mathbf{f}(p)$ is sublinear and lsc, it is the support function of the regular subdifferential. This is the key to a simplified derivation of the regular subdifferential of $\mathbf{f}$. In the next section, we specify suitable inner products for expressing the regular subdifferential using the support function relationship.

4. Inner products

Our derivation of the subdifferential is based on Theorem 3.3 and the relation (2). For this we need to choose inner products on both $\mathcal{P}^n$ and $\mathcal{S}_p$ that are compatible with $\nabla F_p(0)$. The following elementary lemma guides us in these choices. It is the key to both simplifying and clarifying the analysis given in [2]. We leave its proof to the reader. Recall that if L is a linear transformation between the real inner product spaces X and Y, then the adjoint of L, denoted as L^*, is the unique linear transformation from Y to X defined by

$$\left\langle L^* y, x \right\rangle_X = \langle y, Lx \rangle_Y \quad \forall\, y \in Y \text{ and } x \in X.$$

Lemma 4.1. *Let X and Y be finite dimensional vector spaces, and let $L : X \to Y$ be a linear isomorphism.*

(i) *Suppose Y has real inner product $\langle \cdot, \cdot \rangle_Y$ making Y a Euclidean space. Then the bilinear functional $B : X \times X \to \mathbb{R}$ given by $B(x_1, x_2) = \langle Lx_1, Lx_2 \rangle_Y$ is an inner product on X, say $\langle \cdot, \cdot \rangle_{X,L}$. Moreover, the adjoint $L^* : Y \to X$ with respect to the inner products $\langle \cdot, \cdot \rangle_{X,L}$ and $\langle \cdot, \cdot \rangle_Y$ equals L^{-1}.*
(ii) *If X and Y are Euclidean spaces with inner products $\langle \cdot, \cdot \rangle_X$ and $\langle \cdot, \cdot \rangle_Y$, respectively, which satisfy $\langle x_1, x_2 \rangle_X = \langle Lx_1, Lx_2 \rangle_Y$ for all $x_1, x_2 \in X$, then $L^* = L^{-1}$ with respect to these inner products.*

Consider the standard real inner product on $\mathbb{C}^{k+1}$ given by

$$\langle (a_0, a_1, \dots, a_k),\ (b_0, b_1, \dots, b_k)\rangle_{\mathbb{C}^{k+1}} = \sum_{l=0}^{k} \langle a_l,\ b_l\rangle = \sum_{l=0}^{k} \mathrm{Re}(\bar{a}_l b_l),$$

for all $(a_0, a_1, \dots, a_k), (b_0, b_1, \dots, b_k) \in \mathbb{C}^{k+1}$. This inner product induces an inner product on $\mathcal{P}^k$ via Lemma 4.1 with the aid of the *Taylor maps* $\tau_{(k,\lambda_0)} : \mathcal{P}^k \to \mathbb{C}^{k+1}$ defined for each $\lambda_0 \in \mathbb{C}$ and $k = 1, 2, \dots, n$ by

$$\tau_{(k,\lambda_0)}(q) = [q^{(k)}(\lambda_0)/k!, q^{(k-1)}(\lambda_0)/(k-1)!, \dots, q^{(0)}(\lambda_0)], \tag{33}$$

where $q^{(l)}$ denotes the lth derivative of q. The mappings $\tau_{(k,\lambda_0)}$ take a polynomial to its Taylor coefficients at λ_0, and, for each pair (k, λ_0), $\tau_{(k,\lambda_0)}$ is a bijective linear transformation between $\mathcal{P}^k$ and $\mathbb{C}^{k+1}$. Hence, by Lemma 4.1, $\tau_{(k,\lambda_0)}$ induces an inner product on $\mathcal{P}^k$ given by

$$\langle q,\ \tilde{q}\rangle_{(k,\lambda_0)} = \langle \tau_{(k,\lambda_0)}(q),\ \tau_{(k,\lambda_0)}(\tilde{q})\rangle_{\mathbb{C}^{k+1}}, \tag{34}$$

for all $q, \tilde{q} \in \mathcal{P}^k$; moreover, $\tau^*_{(k,\lambda_0)} = \tau^{-1}_{(k,\lambda_0)}$ with respect to these inner products. For future reference, observe that the mapping on $\mathcal{P}^k \times \mathcal{P}^k \times \mathbb{C}$ given by $(q, \tilde{q}, \lambda) \mapsto \langle q,\ \tilde{q}\rangle_{(k,\lambda_0)}$ is continuous since the map $\tilde{\tau}_k : \mathcal{P}^k \times \mathbb{C} \to \mathbb{C}^{k+1}$ given by $\tilde{\tau}_k(q, \lambda) = \tau_{(k,\lambda)}(q)$ is continuous in q and λ [2].

The Taylor maps can be concatenated to build a linear isomorphism between the factorization space $\mathcal{S}_p$ and $\mathbb{C}^{n+1}$ as follows: define $\mathcal{T}_p : \mathcal{S}_p \to \mathbb{C}^{n+1}$ by

$$\begin{aligned}\mathcal{T}_p(u) &= \mathcal{T}_p(\mu_0, u_1, u_2, \dots, u_m)\\ &= [\mu_0, \tau_{(n_1-1,\lambda_1)}(u_1), \tau_{(n_2-1,\lambda_2)}(u_2), \dots, \tau_{(n_m-1,\lambda_m)}(u_m)]\\ &= [\mu_0, (\mu_{11}, \dots, \mu_{1,n_j}), \dots, (\mu_{m1}, \dots, \mu_{mn_m})],\end{aligned} \tag{35}$$

where

$$u = (\mu_0, u_1, u_2, \dots, u_m), \qquad u_j = \sum_{s=1}^{n_j} \mu_{js} e_{(n_j-s,\lambda_j)}, \quad \text{and} \quad \mu_0, \mu_{js} \in \mathbb{C}, \tag{36}$$

for all $s = 1, 2, \dots, n_j$ and $j = 1, 2, \dots, m$. By Lemma 4.1, $\mathcal{T}_p$ induces an inner product $\langle \cdot,\ \cdot\rangle_{\mathcal{S}_p}$ on $\mathcal{S}_p$ by

$$\langle u,\ w\rangle_{\mathcal{S}_p} = \langle \mathcal{T}_p(u),\ \mathcal{T}_p(w)\rangle_{\mathbb{C}^{n+1}}, \tag{37}$$

for all $u, w \in \mathcal{S}_p$, and that with respect to these inner products $\mathcal{T}_p^* = \mathcal{T}_p^{-1}$. It is useful to observe that

$$\langle u,\ w\rangle_{\mathcal{S}_p} = \langle \mathcal{T}_p(u),\ \mathcal{T}_p(w)\rangle_{\mathbb{C}^{n+1}} = \mathrm{Re}[\overline{\mu_0}\omega_0] + \sum_{j=1}^{m} \langle u_j,\ w_j\rangle_{(n_j,\lambda_j)},$$

where u satisfies (36) and, similarly,

$$w = (\omega_0, w_1, w_2, \dots, w_m), \qquad w_j = \sum_{s=1}^{n_j} \omega_{js} e_{(n_j-s,\lambda_j)}, \quad \text{and} \quad \omega_0, \omega_{js} \in \mathbb{C}. \tag{38}$$

Let p be as in (11). We use the mapping $F_p : \mathcal{S}_p \to \mathcal{P}^n$ (13) to construct and inner product on $\mathcal{P}^n$ relative to p. Recall that F_p is a local diffeomorphism at 0, and so the map $\nabla F_p(0) : \mathcal{S}_p \to \mathcal{P}^n$, given by

$$\nabla F_p(0)(q_0, q_1, q_2, \dots, q_m) = q_0 p + \sum_{j=1}^{m} r_j q_j, \tag{39}$$

where $r_j = \prod_{i\neq j} e_{(n_i,\lambda_i)} = p/e_{(n_j,\lambda_j)}$, is an isomorphism. Hence, for every $z, v \in \mathcal{P}^n$, there exists $u \in \mathcal{S}_p$ and $w \in \mathcal{S}_p$ having representations (36) and (38), respectively, such that

$$z = \nabla F_p(0)(\mu_0, u_1, u_2, \dots, u_m) \quad \text{and} \quad v = \nabla F_p(0)(\omega_0, w_1, w_2, \dots, w_m). \tag{40}$$

Again, Lemma 4.1 implies that $\nabla F_p(0)^{-1}$ induces an inner product on $\mathcal{P}^n$ based on the inner product $\langle \cdot,\ \cdot\rangle_{\mathcal{S}_p}$ by setting

$$\begin{aligned}\langle z,\ v\rangle_{(\mathcal{P}^n,p)} &= \langle \nabla F_p(0)^{-1} z,\ \nabla F_p(0)^{-1} v\rangle_{\mathcal{S}_p}\\ &= \langle (\mu_0, u_1, u_2, \dots, u_m),\ (\omega_0, w_1, w_2, \dots, w_m)\rangle_{\mathcal{S}_p}\\ &= \mathrm{Re}(\bar{\mu}_0\omega_0) + \sum_{j=1}^{m}\sum_{s=1}^{n_j} \mathrm{Re}(\bar{\mu}_{js}\omega_{js}),\end{aligned}$$

where z and v are as in (40) and u and w satisfy (36) and (38). Moreover, with respect to these inner products, $\nabla F_p(0)^* = \nabla F_p(0)^{-1}$.

Now consider the composition $\tau_p : \mathcal{P}^n \to \mathbb{C}^{n+1}$ given by

$$\tau_p = \mathcal{T}_p \circ \nabla F_p(0)^{-1}, \tag{41}$$

where $\mathcal{T}_p$ is as in (35) and $\nabla F_p(0)$ is as in (39). For $u, w \in \mathcal{S}_p$ as in (36) and (38) and $z, v \in \mathcal{P}^n$ as in (40), we have

$$\begin{aligned} \langle z,\ v\rangle_{(\mathcal{P}^n,p)} &= \left\langle \nabla F_p(0)^{-1}(z),\ \nabla F_p(0)^{-1}(v)\right\rangle_{\mathcal{S}_p} \\ &= \langle u,\ w\rangle_{\mathcal{S}_p} \\ &= \mathrm{Re}[\overline{\mu_0}\omega_0] + \sum_{j=1}^{m}\sum_{s=1}^{n_j} \mathrm{Re}[\overline{\mu_{js}}\omega_{js}] \\ &= \left\langle \tau_p(z),\ \tau_p(v)\right\rangle_{\mathbb{C}^{n+1}}. \end{aligned} \tag{42}$$

Again by Lemma 4.1, $\tau_p^{-1} = \tau_p^* = \nabla F_p(0) \circ \mathcal{T}_p^{-1}$ with respect to these inner products. The relationship between these spaces is summarized in the diagram below.

$$\begin{tikzcd}
{[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]} \arrow[d, "\nabla F_p(0)"'] \arrow[dr, "\mathcal{T}_p"] & \\
{[\mathcal{P}^n, \langle\cdot,\ \cdot\rangle_{(\mathcal{P}^n,p)}]} \arrow[r, "\tau_p"] & {[\mathbb{C}^{n+1}, \langle\cdot,\ \cdot\rangle_{\mathbb{C}^{n+1}}]}
\end{tikzcd}$$

5. Regular subdifferential and regular normal cone

Theorem 3.1 tells us that $\mathbf{df}(e_{(n,\lambda_0)})$ is proper, lsc, and sublinear under both (A) and (B) when $\partial f(\lambda_0) \neq \{0\}$. Lemma 6 in [9] shows that the expression on the right-hand side of (19) can be written as the support functional for the set

$$\Delta(n, \lambda_0) = \{0\} \times \left(\frac{-1}{n}\partial f(\lambda_0)\right) \times \mathcal{K}(n, \lambda_0) \times \mathbb{C}^{n-2}, \tag{43}$$

where

$$\mathcal{K}(n, \lambda_0) = \begin{cases} \mathcal{K}(0, \lambda_0), & \text{if } f''(\lambda_0) \text{ does not exist,} \\ \{\theta \mid \left\langle \theta,\ f'(\lambda_0)^2\right\rangle \le \left\langle \mathrm{i}f'(\lambda_0),\ f''(\lambda_0)(\mathrm{i}f'(\lambda_0))\right\rangle / n\}, & \text{otherwise,} \end{cases} \tag{44}$$

with

$$\mathcal{K}(0, \lambda_0) = -\mathrm{cone}(\partial f(\lambda_0)^2) + \mathrm{i}\left[\mathrm{rspan}\left(\partial f(\lambda_0)^2\right)\right].$$

That is,

$$\sigma_{\Delta(n,\lambda_0)}(w) = \begin{cases} (f'(\lambda_0; \omega_1) + f''(\lambda_0; \sqrt{-\omega_2}, \sqrt{-\omega_2}))/n, & \text{if (16) and (17) hold,} \\ +\infty, & \text{otherwise,} \end{cases} \tag{45}$$

where we use the convention (20) when (B) holds at λ_0. This gives the following characterization of the regular subdifferential in the one root case.

Theorem 5.1 ([9, Theorem 8]). *Let $\lambda_0 \in \mathrm{dom}(\partial f)$ be such that $\partial f(\lambda_0) \neq \{0\}$. Then, relative to the inner product $\langle\cdot,\ \cdot\rangle_{(n,\lambda_0)}$ in (34),*

$$\hat{\partial}\mathbf{f}(e_{(n,\lambda_0)}) \supset \tau^*_{(n,\lambda_0)}(\Delta_0(n, \lambda_0)),$$

where

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

and $\tau_{(n,\lambda_0)}$ is defined in (33). If either (A) or (B) holds at $\lambda = \lambda_0$, then

$$\mathbf{df}(e_{(n,\lambda_0)})(v) = \sigma_{\hat{\partial}\mathbf{f}(e_{(n,\lambda_0)})}(v) \quad \forall\, v \in \mathcal{P}^n, \text{ with } \hat{\partial}\mathbf{f}(e_{(n,\lambda_0)}) = \tau^*_{(n,\lambda_0)}(\Delta(n, \lambda_0)),$$

and

$$\mathbf{df}_1(e_{(n,\lambda_0)}) = \sigma_{\hat{\partial}\mathbf{f}_1(e_{(n,\lambda_0)})}(v) \quad \forall\, v \in \mathcal{P}^n, \text{ with } \hat{\partial}\mathbf{f}_1(e_{(n,\lambda_0)}) = \tau^*_{(n,\lambda_0)}(\Delta_1(n, \lambda_0)),$$

where

$$\Delta_1(n,\lambda_0)=\mathbb{C}\times\left(\frac{-1}{n}\partial f(\lambda_0)\right)\times\mathcal{K}(n,\lambda_0)\times\mathbb{C}^{n-2}.$$

Let $v=\nabla F_p(0)(\omega_0,w_1,\ldots,w_m)$ for $(\omega_0,w_1,\ldots,w_m)\in\mathcal{S}_p$. Recall from Theorem 3.3 that if $p\in\mathcal{M}^n$ has prime factorization (11), then

$$\mathbf{df}(p)(v)\geq\max_{j\in\mathcal{I}(p)}\{f'(\lambda_j;-\omega_{j1})/n_j\}$$

whenever $(\omega_0,w_1,\ldots,w_m)$ satisfies (29) and (30) for all $j\in\mathcal{I}(p)$; otherwise, $\mathbf{df}(p)(v)=+\infty$. If, in addition, f satisfies either (A) or (B) at $\lambda=\lambda_j$ for all $j\in\mathcal{I}(p)$, then

$$\mathbf{df}(p)(v)=\max_{j\in\mathcal{I}(p)}\{[f'(\lambda_j;-\omega_{j1})+f''(\lambda_j;\sqrt{-\omega_{j2}},\sqrt{-\omega_{j2}})]/n_j\},\tag{46}$$

whenever $(\omega_0,w_1,\ldots,w_m)$ satisfies (29) and (30) for all $j\in\mathcal{I}(p)$, where we use the convention (20) if (B) holds. By Theorem 5.1, each term appearing in the maximum in the right-hand side of (46) is $\mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)})(w_j)$, where $\mathbf{f}_{[1,n_j]}:\mathcal{P}^{n_j}\to\overline{\mathbb{R}}$ is defined in (21). Therefore, if f satisfies either (A) or (B) at λ_j for each λ_j with $j\in\mathcal{I}(p)$, then

$$\mathbf{df}(p)(v)=\max_{j\in\mathcal{I}(p)}\mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)})(w_j)=\max_{j\in\mathcal{I}(p)}\sigma(w_j\mid\hat{\partial}\mathbf{f}_{[1,n_j]}(e_{(n_j,\lambda_j)})),\tag{47}$$

where we think of each w_j as an element of $\mathcal{P}^{n_j,(n_j-1)}$ rather than $\mathcal{P}^{n_j-1}$ so that the domain requirements for $\mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)})$ are satisfied (note that $\mathrm{dom}(\mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)}))\subset\mathcal{P}^{n_j,(n_j-1)}$). Moreover, again by Theorem 5.1,

$$\hat{\partial}\mathbf{f}_{[1,n_j]}(e_{(n_j,\lambda_j)})=\tau^*_{(n_j,\lambda_j)}(\Delta_1(n_j,\lambda_j)).$$

Therefore, it seems that a formula for the subdifferential of $\mathbf{f}$ at p can be obtained as a straightforward consequence the following elementary fact from convex analysis.

Proposition 5.2 *([13, Theorem C.3.3.2(ii)]). Let E be a Euclidean space and $\mathcal{I}$ an arbitrary index set. Let $C^i\subset E$ be closed and nonempty for all $i\in\mathcal{I}$. Then*

$$\max_{i\in\mathcal{I}}\sigma_{C^i}(v)=\sigma(v\mid\mathrm{conv}(\cup_{i\in\mathcal{I}}C^i))$$

for all $v\in E$.

However, (47) is deficient since the argument on the left-hand side is v, whereas the argument in each term in the maximum on the right-hand side is w_j. We correct this problem by slightly modifying the definitions of the sets $\Delta_0(n_j,\lambda_j)$ and $\Delta(n_j,\lambda_j)$, and then extending them to $\mathcal{S}_p$. Let $\mathcal{K}(n_j,\lambda_j)$ be as in (43), τ_p as in (41), and set

$$\hat{\Delta}(n_j,\lambda_j)=\left(\frac{-1}{n_j}\partial f(\lambda_j)\right)\times\mathcal{K}(n_j,\lambda_j)\times\mathbb{C}^{n_j-2},$$
$$\hat{\Delta}_0(n_j,\lambda_j)=\left(\frac{-1}{n_j}\partial f(\lambda_j)\right)\times\mathcal{K}(0,\lambda_j)\times\mathbb{C}^{n_j-2},$$
$$D^j=\mathcal{T}_p^*(0,\ldots,0,\hat{\Delta}(n_j,\lambda_j),0,\ldots,0),\quad\text{and}$$
$$D_0^j=\mathcal{T}_p^*(0,\ldots,0,\hat{\Delta}_0(n_j,\lambda_j),0,\ldots,0),$$

where in both D^j and D_0^j, for $j=1,\ldots,m$, the nonzero entries occur in the jth component with the component indexing starting from zero so that the first component is always the scalar zero. The sets D^j and D_0^j all lie in $\mathcal{S}_p$. Finally, set

$$D(p)=\mathrm{conv}\cup_{j\in\mathcal{I}(p)}D^j,\quad\text{and}\quad D_0(p)=\mathrm{conv}\cup_{j\in\mathcal{I}(p)}D_0^j.$$

Let $\langle\cdot,\cdot\rangle_{\mathcal{P}^n}$ be a given inner product on $\mathcal{P}^n$ (not necessarily the inner product $\langle\cdot,\cdot\rangle_{(\mathcal{P}^n,p)}$), and let $\nabla F_p(0)^*$ denote the adjoint of $\nabla F_p(0)$ with respect to the inner products $\langle\cdot,\cdot\rangle_{\mathcal{P}^n}$ and $\langle\cdot,\cdot\rangle_{\mathcal{S}_p}$.

Theorem 5.3. *Let $f:\mathbb{C}\to\overline{\mathbb{R}}$ be proper, convex and lsc, and let $p\in\mathrm{dom}(\mathbf{f})\cap\mathcal{M}_1^n$ have prime factorization (11) where $\partial f(\lambda_j)\neq\{0\}$ for each $j\in\mathcal{I}(p)$. Then, with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathcal{P}^n}$,*

$$\hat{\partial}\mathbf{f}(p)\supset\{z\mid\nabla F_p(0)^*z\in D_0(p)\}.\tag{48}$$

If f satisfies either (A) or (B) at λ_j, for each $j\in\mathcal{I}(p)$, then

$$\hat{\partial}\mathbf{f}(p)=\{z\mid\nabla F_p(0)^*z\in D(p)\}\quad\text{and}\quad\mathbf{df}(p)(v)=\sigma_{\hat{\partial}\mathbf{f}(p)}(v).\tag{49}$$

In particular, if $\langle\cdot,\cdot\rangle_{\mathcal{P}^n}$ is chosen to be $\langle\cdot,\cdot\rangle_{(\mathcal{P}^n,p)}$, then $\hat{\partial}\mathbf{f}(p)=\nabla F_p(0)D(p)$.

Remark 3. The representation $\hat{\partial}\mathbf{f}(p) = \nabla F_p(0)D(p)$ under inner product $\langle\cdot,\ \cdot\rangle_{(\mathcal{P}^n,p)}$ is new even in the case of the polynomial abscissa mapping. It can be used to simplify the representation of the regular subdifferential for the abscissa mapping given in [2, Theorem 2.2]. We use it in the final section of this paper to represent the subdifferential of the polynomial radius mapping $\mathbf{r}$.

Proof. Let $v \in \mathcal{P}^n$ and $(\omega_0, w_1, \dots, w_m) \in \mathcal{S}_p$ be such that

$$v = \nabla F_p(0)(\omega_0, w_1, \dots, w_m). \tag{50}$$

The proof of (48) is nearly identical to that of (49) if one uses Theorem 3.3 and (45) to observe that

$$\begin{aligned}\mathbf{df}(p)(v) \geq \max_{j\in\mathcal{I}(p)}\{f'(\lambda_j; -\omega_{j1})/n_j\} &= \max_{j\in\mathcal{I}(p)} \sigma(\tau_{(n_j-1,\lambda_j)}(w_j) \mid \hat{\Delta}_0(n_j, \lambda_j))\\ &= \max_{j\in\mathcal{I}(p)} \sigma(w_j \mid \tau^*_{(n_j-1,\lambda_j)}(\hat{\Delta}_0(n_j, \lambda_j))).\end{aligned}$$

Therefore, we only prove (49). Suppose that f satisfies either (A) or (B) at λ_j, for all $j \in \mathcal{I}(p)$. Since each of the sets D^j is closed, convex and nonempty, the result will follow from (47) and Proposition 5.2 if we show that $\sigma_{D^j}(v) = \mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)})(w_j)$, where v and w satisfy (50).

First note that for each $d_j \in D^j$ there exists $u_j \in \hat{\Delta}(n_j, \lambda_j)$ such that

$$\nabla F_p(0)^* d_j = \mathcal{T}_p^*(0, \dots, 0, u_j, 0, \dots, 0) = (0, \dots, 0, \tau^*_{(n_j-1,\lambda_j)}(u_j), 0, \dots, 0).$$

Let $v \in \mathcal{P}^n$ and $w = (\omega_0, w_1, \dots, w_m) \in \mathcal{S}_p$ be as in (50). Then

$$\begin{aligned}\sigma_{D^j}(v) &= \sup_{\nabla F_p(0)^* d_j\in D^j} \langle d_j,\ v\rangle_{\mathcal{P}^n} = \sup_{\nabla F_p(0)^* d_j\in D^j} \langle d_j,\ \nabla F_p(0)w\rangle_{\mathcal{P}^n}\\ &= \sup_{\nabla F_p(0)^* d_j\in D^j} \left\langle \nabla F_p(0)^* d_j,\ (\omega_0, w_1, \dots, w_m)\right\rangle_{\mathcal{S}_p}\\ &= \sup_{u_j\in\hat{\Delta}(n_j,\lambda_j)} \left\langle \mathcal{T}_p^*(0, \dots, 0, u_j, 0, \dots, 0),\ (\omega_0, w_1, \dots, w_m)\right\rangle_{\mathcal{S}_p}\\ &= \sup_{u_j\in\hat{\Delta}(n_j,\lambda_j)} \left\langle \tau^*_{(n_j-1,\lambda_j)}(u_j),\ w_j\right\rangle_{(n_j-1,\lambda_j)}\\ &= \sup_{u_j\in\hat{\Delta}(n_j,\lambda_j)} \left\langle u_j,\ \tau_{(n_j-1,\lambda_j)}(w_j)\right\rangle_{\mathbb{C}^{n_j}} = \sup_{u_j\in\Delta(n_j,\lambda_j)} \left\langle u_j,\ \tau_{(n_j,\lambda_j)}(w_j)\right\rangle_{\mathbb{C}^{n_j+1}}\\ &= \sup_{u_j\in\Delta(n_j,\lambda_j)} \left\langle \tau^*_{(n_j,\lambda_j)}(u_j),\ w_j\right\rangle_{(n_j,\lambda_j)} = \sup_{r_j\in\tau^*_{(n_j,\lambda_j)}(\Delta(n_j,\lambda_j))} \langle r_j,\ w_j\rangle_{(n_j,\lambda_j)}\\ &= \sigma(w_j \mid \tau^*_{(n_j,\lambda_j)}(\Delta(n_j, \lambda_j))) = \mathbf{df}_{[1,n_j]}(e_{(n_j,\lambda_j)})(w_j),\end{aligned}$$

where the fifth line follows since the first component of $\Delta(n_j, \lambda_j)$ is zero (see (43)).

The final statement of the theorem follows since $\nabla F_p(0)^* = \nabla F_p(0)^{-1}$ when the inner product on $\mathcal{P}^n$ is given by $\langle\cdot,\ \cdot\rangle_{(\mathcal{P}^n,p)}$. $\square$

The formulas for the subderivative and subdifferential for any $p \in \mathrm{dom}(\mathbf{f}) \cap \mathcal{M}^n$ can be obtained by applying the following elementary lemma.

Lemma 5.4. *Let $\mathbf{h} : \mathcal{P}^n \to \overline{\mathbb{R}}$ be a weak prf. Given $p \in \mathrm{dom}(\mathbf{h})$ and $\kappa \in \mathbb{C} \setminus \{0\}$, we have*

$$\kappa p \in \mathrm{dom}(\mathbf{h}) \quad \textit{and} \quad \mathbf{dh}(\kappa p)(v) = \mathbf{dh}(p)(\kappa^{-1}v) \quad \forall v \in \mathcal{P}^n.$$

Moreover, if $\mathbf{dh}(p) = \sigma_{\hat{\partial}\mathbf{h}(p)}$, then $\hat{\partial}\mathbf{h}(\kappa p) = \bar{\kappa}^{-1}\hat{\partial}\mathbf{h}(p)$.

Proof. The domain property follows immediately from the definition of a weak prf. The subderivative equivalence follows from the definitions of weak prf and the subderivative. The final equivalence follows since

$$\mathbf{dh}(\kappa p)(v) = \mathbf{dh}(p)(\kappa^{-1}v) = \sup_{u\in\hat{\partial}\mathbf{h}(p)} \left\langle u,\ \kappa^{-1}v\right\rangle_{\mathcal{P}^n} = \sup_{u\in\bar{\kappa}^{-1}\hat{\partial}\mathbf{h}(p)} \langle u,\ v\rangle_{\mathcal{P}^n}. \quad \square$$

Given $p \in \mathcal{P}^n$ having factorization (11), Theorem 5.3 in conjunction with the relationship (3) can be used to obtain a representation for the regular normal cone to epi($\mathbf{f}$). The only obstacle to this being a straightforward computation is the absence of a formula for the recession cone $\{z \mid \nabla F_p(0)^* z \in D(p)\}^\infty$. This is provided in the following proposition.

Proposition 5.5. *Let p, f, D, and $\langle\cdot,\ \cdot\rangle_{\mathcal{P}^n}$ be as in the statement of Theorem 5.3. Then*

$$\{z \mid \nabla F_p(0)^*z \in D(p)\}^\infty = \{z \mid \nabla F_p(0)^*z \in \operatorname{conv}\cup_{j\in\mathcal{I}(p)}(D^j)^\infty\},$$

where $(D^j)^\infty = \mathcal{T}_p^*(0,\ldots,0,\hat{\Delta}(n_j,\lambda_j)^\infty,0,\ldots,0)$ *and*

$$\hat{\Delta}(n_j,\lambda_j)^\infty = \frac{-1}{n_j}\partial f(\lambda_j)^\infty \times \mathcal{K}(0,\lambda_j)\times\mathbb{C}^{n_j-2}.$$

That is, $z \in \{w \mid \nabla F_p(0)^*w \in D(p)\}^\infty$ *if and only if there exists a point* $(\mu_0,u_1,\ldots,u_m)\in\mathcal{S}_p$ *such that* $\nabla F_p(0)^*(z) = (\mu_0,u_1,\ldots,u_m)$, *with* $u_j = \sum_{s=1}^{n_j}\mu_{js}e_{(n_j-s,\lambda_j)}$ *for* $j=1,\ldots,m$, *and* $\mu_0,\mu_{js}\in\mathbb{C}$ *for* $s=1,\ldots,n_j,\ j=1,\ldots,m$, *satisfying*

$$\mu_0 = 0,\qquad u_j = 0\quad \text{for } j\notin\mathcal{I}(p),\ \text{and}$$
$$\mu_{j1}\in -\partial f(\lambda_j)^\infty\quad\text{and}\quad \mu_{j2}\in\mathcal{K}(0,\lambda_j)\ \forall j\in\mathcal{I}(p).$$

Proof. Well-known properties of the horizon cone give

$$\{z \mid \nabla F_p(0)^*z\in D(p)\}^\infty = \{z \mid \nabla F_p(0)^*z \in D(p)^\infty\}\quad [12, \text{Proposition 2.1.11}]$$

as well as

$$\begin{aligned}
D(p)^\infty &= \operatorname{conv}\left[\left(\cup_{j\in\mathcal{I}(p)}(D^j)\right)^\infty\right] \quad [12, \text{Lemma 2.3.2}]\\
&= \operatorname{conv}\cup_{j\in\mathcal{I}(p)}\left[(D^j)^\infty\right]\quad [12, \text{Proposition 2.1.9}]\\
&= \operatorname{conv}\cup_{j\in\mathcal{I}(p)}\mathcal{T}_p^*[(0,\ldots,0,\hat{\Delta}(n_j,\lambda_j),0,\ldots,0)^\infty]\quad [12, \text{Corollary 2.3.2}]\\
&= \operatorname{conv}\cup_{j\in\mathcal{I}(p)}\mathcal{T}_p^*(0,\ldots,0,\hat{\Delta}(n_j,\lambda_j)^\infty,0,\ldots,0)\quad [12, \text{Proposition 2.1.10}],
\end{aligned}$$

with

$$\hat{\Delta}(n_j,\lambda_j)^\infty = \left(\frac{-1}{n_j}\partial f(\lambda_j)^\infty\right)\times\mathcal{K}(n_j,\lambda_j)^\infty\times\mathbb{C}^{n_j-2}\quad [12, \text{Proposition 2.1.10}].$$

Finally, the equivalence $\mathcal{K}(n_j,\lambda_j)^\infty = \mathcal{K}(0,\lambda_j)$ follows from [9, Equation (49)] which proves the result. □

6. Subdifferential regularity

The derivation of formulas for general and horizon subgradients of $\mathbf{f}$ at a polynomial p requires taking limits of regular subgradients $g^\nu \in \hat{\partial}\mathbf{f}(p^\nu)$ where $p^\nu \xrightarrow{\nu} p$. The formulas for regular subgradients given in Theorem 5.3 depend on the factorization space $\mathcal{S}_p$ and the choice of an inner product $\langle\cdot,\ \cdot\rangle_{\mathcal{P}^n}$ on $\mathcal{P}^n$. Therefore, the limiting behavior of regular subgradients is tied to the limiting behavior of the factorization spaces $\mathcal{S}_p$ as well as the mappings $\nabla F_{p^\nu}(0)^*$ and $\mathcal{T}_{p^\nu}$ along sequences $\{p^\nu\}$ converging to p. One of the difficulties associated with these limits is that although the mappings $\nabla F_{p^\nu}(0)$ are invertible for each ν, the limit of these transformations is typically not invertible. Much of the machinery we use to handle these kinds of sequences is developed in [2]. We first review this material and then augment it with ideas from Section 4. We begin with the spaces $\mathcal{S}_{p^\nu}$.

Let $p\in\mathcal{M}^n\cap\operatorname{dom}(\mathbf{f})$ have factorization (11), and consider $p^\nu\xrightarrow{\nu}p$ with $\{p^\nu\}\subset\operatorname{dom}(\mathbf{f})\cap\mathcal{M}^n$. By Lemma 2.1, we can assume $\{p^\nu\}\subset\mathcal{M}_1^n$. Moreover, since $p^\nu\xrightarrow{\nu}p$, Lemma 1.4 in [2] tells us that we can write

$$p^\nu = \prod_{j=1}^m q_j^\nu\quad\text{and}\quad q_j^\nu = \prod_{s=1}^{l_j^\nu} e_{(n_{js}^\nu,\lambda_{js}^\nu)},$$

where

$$\deg(q_j^\nu) = n_j,\qquad q_j^\nu\xrightarrow{\nu}e_{(n_j,\lambda_j)},\qquad \sum_{s=1}^{l_j^\nu}n_{js}^\nu = n_j,\qquad \lambda_{js}^\nu\xrightarrow{\nu}\lambda_{js},\qquad \lambda_{js}=\lambda_j,$$

and $\lambda_{js}^\nu\neq\lambda_{it}^\nu$ if either $j=i$ and $s\neq t$ or $j\neq i$, for $s=1,\ldots,l_j$ and $j=1,\ldots,m$. Since there are only finitely many partitions of n, by going to a subsequence if necessary, we can assume that

$$l_j^\nu = l_j\quad\text{and}\quad n_{js}^\nu = n_{js}\quad\text{for all } \nu=1,2,\ldots.$$

Define the factorization space $\tilde{\mathcal{S}}$ by

$$\tilde{\mathcal{S}} = \mathbb{C}\times\hat{\mathcal{S}}_{\pi_1}\times\hat{\mathcal{S}}_{\pi_1}\times\cdots\times\hat{\mathcal{S}}_{\pi_m},$$

where

$$\hat{\mathcal{S}}_{\pi_j} = \mathcal{P}^{n_{j1}-1}\times\cdots\times\mathcal{P}^{n_{jl_j}-1}\quad\text{and}\quad \pi_j=(n_{j1},\ldots,n_{jl_j})\quad j=1,\ldots,m.$$

The factorization spaces $\mathcal{S}_{p^\nu}$ and $\tilde{\mathcal{S}}$ coincide up to a permutation of the components, for all $\nu = 1, 2, \ldots$. We suppress this permutation and simply write $\mathcal{S}_{p^\nu} = \tilde{\mathcal{S}}$ for all $\nu = 1, 2, \ldots$.

Next consider the mappings $\nabla F_{p^\nu}(0) : \tilde{\mathcal{S}} \to \mathcal{P}^n$ given by

$$\nabla F_{p^\nu}(0)(\omega_0, (w_{11}, \ldots, w_{1l_1}), \ldots, (w_{m1}, \ldots w_{ml_m})) = r_0^\nu \omega_0 + \sum_{j=1}^{m} r_j^\nu \left(\sum_{s=1}^{l_j} \hat{r}_{js}^\nu w_{js} \right) \tag{51}$$

where

$$r_j^\nu = \prod_{i \neq j} \prod_{s=1}^{l_i} e_{(n_{is}, \lambda_{is}^\nu)} = p^\nu / q_j^\nu, \qquad \hat{r}_{js}^\nu = \prod_{t \neq s}^{l_j} e_{(n_{jt}, \lambda_{jt}^\nu)} = q_j^\nu / e_{(n_{js}, \lambda_{js}^\nu)}, \quad \text{and} \quad r_0^\nu = p^\nu.$$

The representation (51) shows that the mappings $\nabla F_{p^\nu}(0)$ can be written in factored form as

$$\nabla F_{p^\nu}(0) = \Gamma_\nu \circ \Psi_\nu,$$

where $\Gamma_\nu : \mathcal{S}_p \to \mathcal{P}^n$ and $\Psi_\nu : \tilde{\mathcal{S}} \to \mathcal{S}_p$ are given by

$$\Gamma_\nu(\mu_0, u_1, \ldots, u_m) = r_0^\nu \mu_0 + \sum_{j=1}^{m} r_j^\nu u_j \quad \text{and} \quad \Psi_\nu = [I, \psi_{\nu,1}, \ldots, \psi_{\nu,m}],$$

with $\psi_{\nu,j} : \hat{\mathcal{S}}_{\pi_j} \to \mathcal{P}^{n_j - 1}$ given by

$$\psi_{\nu,j}(w_{j1}, \ldots, w_{jl_j}) = \sum_{s=1}^{l_j} \hat{r}_{js}^\nu w_{js} \quad j = 1, \ldots, m.$$

These mappings have well-defined limits as $\nu \to \infty$. Indeed, if, for $j = 1, \ldots, m$, we define the mappings $\psi_{(\pi_j, \lambda_j)} : \hat{\mathcal{S}}_{\pi_j} \mapsto \mathcal{P}^{n_j - 1}$ by

$$\psi_{(\pi_j, \lambda_j)}(a_{j1}, a_{j2}, \ldots, a_{jl_j}) = \sum_{s=1}^{l_j} e_{(n_j - n_{js}, \lambda_j)} a_{js},$$

then

$$\Gamma_\nu \xrightarrow{\nu} \nabla F_p(0), \quad \Psi_\nu \xrightarrow{\nu} \Psi = [I, \psi_{(\pi_1, \lambda_1)}, \psi_{(\pi_2, \lambda_2)}, \ldots, \psi_{(\pi_m, \lambda_m)}], \quad \text{and}$$
$$\nabla F_{p^\nu}(0) \xrightarrow{\nu} \nabla F_p(0) \circ \Psi = \Xi,$$

where convergence is with respect to any choice of norms on $\mathcal{P}^n$, $\mathcal{S}_p$, and $\tilde{\mathcal{S}}$. The operators also allow us to compute limits of the operators $\nabla F_p(0)^*$ which is necessary for computing the limits of regular subgradients. For this we will again need a suitable choice of inner products on the various spaces. The following lemma provides the key.

Lemma 6.1 *([2, Lemma 3.1]). For each $j = 1, \ldots, m$, the inner products*

$$\left\langle (u_{j1}, \ldots, u_{jl_j}),\ (w_{j1}, \ldots, w_{jl_j}) \right\rangle_{(\nu, \hat{\mathcal{S}}_{\pi_j})} = \sum_{s=1}^{l_j} \left\langle u_{js},\ w_{js} \right\rangle_{(n_{js}-1, \lambda_{js}^\nu)}$$

converge pointwise to the inner product

$$\left\langle (u_{j1}, \ldots, u_{jl_j}),\ (w_{j1}, \ldots, w_{jl_j}) \right\rangle_{(\infty, \hat{\mathcal{S}}_{\pi_j})} = \sum_{s=1}^{l_j} \left\langle u_{js},\ w_{js} \right\rangle_{(n_{js}-1, \lambda_j)}.$$

Moreover, for each $j = 1, \ldots, m$, the adjoint transformations $\psi_{\nu,j}^* : \mathcal{P}^{n_j-1} \to \hat{\mathcal{S}}_{\pi_j}$, with respect to the Euclidean spaces $[\hat{\mathcal{S}}_{\pi_j}, \langle \cdot, \cdot \rangle_{(\nu, \hat{\mathcal{S}}_{\pi_j})}]$ and $[\mathcal{P}^{n_j-1}, \langle \cdot, \cdot \rangle_{(n_j-1, \lambda_j)}]$, converge to the adjoint $\psi_{(\pi_j, \lambda_j)}^*$ with respect to the Euclidean spaces $[\hat{\mathcal{S}}_{\pi_j}, \langle \cdot, \cdot \rangle_{(\infty, \hat{\mathcal{S}}_{\pi_j})}]$ and $[\mathcal{P}^{n_j-1}, \langle \cdot, \cdot \rangle_{(n_j-1, \lambda_j)}]$ with

$$\psi_{(\pi_j, \lambda_j)}^* \left(\sum_{s=1}^{n_j} \beta_s e_{(n_j - s, \lambda_j)} \right) = \left[\sum_{s=1}^{n_{j1}} \beta_s e_{(n_{j1} - s, \lambda_j)}, \ldots, \sum_{s=1}^{n_{jl_j}} \beta_s e_{(n_{jl_j} - s, \lambda_j)} \right].$$

Proof. The convergence of the inner products follows immediately from the continuity of the mapping $\tilde{\tau}_k : \mathcal{P}^k \times \mathbb{C} \to \mathbb{C}^{k+1}$ given by $\tilde{\tau}_k(q, \lambda) = \tau_{(k,\lambda)}(q)$. The convergence of the adjoints follows from the convergence of the inner products and the definition of the adjoint. The representation for $\psi_{(\pi_j, \lambda_j)}^*$ is proved in [2, Lemma 3.1]. $\square$

Since $\Gamma_\nu \xrightarrow{\nu} \nabla F_p(0)$, we have that $\Gamma_\nu^* \xrightarrow{\nu} \nabla F_p(0)^*$, where all of these adjoints are taken with respect to the Euclidean spaces $[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]$ and $[\mathcal{P}^n, \langle\cdot,\ \cdot\rangle_{\mathcal{P}^n}]$, where $\langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}$ is defined in (37). It is important that these inner products are fixed and do not change with ν. Lemma 6.1 implies that $\Psi_\nu^* \xrightarrow{\nu} \Psi^*$ where Ψ^* is the adjoint with respect to the Euclidean spaces $[\tilde{\mathcal{S}}, \langle\cdot,\ \cdot\rangle_{(\infty,\tilde{\mathcal{S}})}]$ and $[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]$ with

$$\langle u,\ w\rangle_{(\infty,\tilde{\mathcal{S}})} = \mathrm{Re}(\bar{\mu}_0\omega_0) + \sum_{j=1}^{m} \left\langle u_j,\ w_j\right\rangle_{(\infty,\hat{\mathcal{S}}_{\pi_j})},$$

and each Ψ_ν^* is the adjoint with respect to the Euclidean spaces $[\tilde{\mathcal{S}}, \langle\cdot,\ \cdot\rangle_{(\nu,\tilde{\mathcal{S}})}]$ and $[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]$ with

$$\langle u,\ w\rangle_{(\nu,\tilde{\mathcal{S}})} = \mathrm{Re}(\bar{\mu}_0\omega_0) + \sum_{j=1}^{m} \left\langle u_j,\ w_j\right\rangle_{(\nu,\hat{\mathcal{S}}_{\pi_j})}.$$

Therefore $\nabla F_{p^\nu}(0)^* \xrightarrow{\nu} \Psi^* \circ \nabla F_p(0)^* = \Xi^*$.

We summarize the relationships between the mappings and inner product spaces in the diagram below.

$$\begin{tikzcd}
{[\tilde{\mathcal{S}}, \langle\cdot,\ \cdot\rangle_{(\nu,\tilde{\mathcal{S}})}]} \arrow[r, "\Psi_\nu"] \arrow[d, "\nabla F_{p^\nu}(0)"'] & {[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]} \arrow[dl, "\Gamma_\nu"] & {} \arrow[r, "\nu\to\infty"] & {} & {[\tilde{\mathcal{S}}, \langle\cdot,\ \cdot\rangle_{(\infty,\tilde{\mathcal{S}})}]} \arrow[r, "\Psi"] \arrow[d, "\Xi"'] & {[\mathcal{S}_p, \langle\cdot,\ \cdot\rangle_{\mathcal{S}_p}]} \arrow[dl, "\nabla F_p(0)"] \\
{[\mathcal{P}^n, \langle\cdot,\ \cdot\rangle_{\mathcal{P}^n}]} & & & & {[\mathcal{P}^n, \langle\cdot,\ \cdot\rangle_{\mathcal{P}^n}]} &
\end{tikzcd}$$

We are now ready to establish the subdifferential regularity of $\mathbf{f}$.

Theorem 6.2. *Let $f : \mathbb{C} \to \mathbb{R}$ be convex and let $p \in \mathrm{dom}(\mathbf{f}) \cap \mathcal{M}_1^n$ as in (11) be such that f is twice continuously differentiable at λ_j with $f'(\lambda_j) \neq 0$ and satisfying (A) at $\lambda = \lambda_j$ for all $j \in \mathcal{I}(p)$. Then $\mathbf{f}$ is subdifferentially regular at p, that is, $\partial\mathbf{f}(p) = \hat{\partial}\mathbf{f}(p)$ and $\partial^\infty\mathbf{f}(p) = \hat{\partial}\mathbf{f}(p)^\infty$.*

Proof. It is always the case that $\partial\mathbf{f}(p) \supset \hat{\partial}\mathbf{f}(p)$ and $\partial^\infty\mathbf{f}(p) \supset \hat{\partial}\mathbf{f}(p)^\infty$, so we need only show the reverse inclusions. The proofs in both cases are nearly identical and so we only provide a proof for the somewhat more difficult inclusion $\partial\mathbf{f}(p) \subset \hat{\partial}\mathbf{f}(p)$.

Since f is twice continuously differentiable at each λ_j for $j \in \mathcal{I}(p)$, there is a neighborhood U of p such that for all $q \in U$ we have $\mathcal{I}(q) \subset \mathcal{I}(p)$ and for all $\lambda \in \mathcal{R}(q)$ with $f(\lambda) = \mathbf{f}(q)$ it must be the case that $f'(\lambda) \neq 0$ and (A) is satisfied at λ. Therefore, on U, the regular subdifferential of $\mathbf{f}$ is given by Theorem 5.3. Let $p^\nu \xrightarrow{\nu} p$ and $z^\nu \xrightarrow{\nu} z$ with $z^\nu \in \hat{\partial}\mathbf{f}(p^\nu)$ for all $\nu = 1, 2, \ldots$. We need to show that $z \in \hat{\partial}\mathbf{f}(p)$. With no loss in generality, we can assume that $\{p^\nu\} \subset U$ so that $\hat{\partial}\mathbf{f}(p^\nu)$ is given by Theorem 5.3 for all $\nu = 1, 2, \ldots$. By Lemma 5.4, we can also assume with no loss in generality that $\{p^\nu\} \subset \mathcal{M}_1^n$. Set

$$w^\nu = \Gamma_\nu^* z^\nu \quad \text{and} \quad u^\nu = \Psi_\nu^* w^\nu \quad \text{so that } u^\nu = \Psi_\nu^* \Gamma_\nu^* z^\nu, \tag{52}$$

and, using Theorem 5.3, for all $\nu = 1, 2, \ldots$, write

$$w^\nu = (\omega_0^\nu, w_1^\nu, \ldots, w_m^\nu) \in \mathcal{S}_p \quad \text{with } w_j^\nu \in \mathcal{P}^{n_j - 1},\ 1 \le j \le m, \tag{53}$$

$$u^\nu = (\mu_0^\nu, u_1^\nu, \ldots, u_m^\nu) \in \tilde{\mathcal{S}}, \tag{54}$$

$$u_j^\nu = (u_{j1}^\nu, \ldots, u_{jl_j}^\nu) \in \hat{\mathcal{S}}_{\pi_j}, \quad 1 \le j \le m,\ \text{with} \tag{55}$$

$$u_{js}^\nu = \gamma_{js}^\nu \sum_{t=1}^{n_{js}} \mu_{jst}^\nu e_{(n_{js}-t,\lambda_{js}^\nu)} \in \mathcal{P}^{n_{js}-1}, \quad 1 \le j \le m,\ 1 \le s \le l_j, \tag{56}$$

where, for $1 \le j \le m,\ 1 \le s \le l_j$,

$$\gamma_{js}^\nu \ge 0 \quad \text{with } \gamma_{js}^\nu = 0 \text{ if } (j,s) \notin \mathcal{I}_\nu = \{(j,s) \mid \mathbf{f}(p^\nu) = f(\lambda_{js}^\nu)\}, \qquad \sum_{j=1}^{m} \sum_{s=1}^{l_j} \gamma_{js}^\nu = 1, \tag{57}$$

and, for $(j,s) \in \mathcal{I}_\nu$,

$$\mu_{js1}^\nu = \frac{1}{n_{js}} f'(\lambda_{js}^\nu), \qquad \mu_{js2}^\nu \in \mathcal{K}(n_{js}, \lambda_{js}^\nu), \quad \text{and} \quad \mu_{jst}^\nu \in \mathbb{C}, \quad 3 \le t \le n_{js}. \tag{58}$$

The continuity of the roots of a polynomial (including multiplicities) on $\mathcal{M}^n$ implies that $\lambda_{js}^\nu \xrightarrow{\nu} \lambda_j$ for $1 \le j \le m$ and $1 \le s \le n_{js}$. Since there are only finitely many possibilities for the index set $\mathcal{I}_\nu$, we can assume with no loss in generality that there is an index set $\tilde{\mathcal{I}}$ such that $\mathcal{I}_\nu = \tilde{\mathcal{I}}$ for all $\nu = 1, 2, \ldots$. Moreover, by continuity, it must be the case that $\{j \mid \exists s \text{ such that } (j, s) \in \tilde{\mathcal{I}}\} \subset \mathcal{I}(p)$. Since $\tilde{\mathcal{I}}$ is fixed, the compactness of the set of possible γ_{js}^ν's implies that we can also assume with no loss in generality that there exist γ_{js} such that $\gamma_{js}^\nu \xrightarrow{\nu} \gamma_{js}$ for $1 \le j \le m,\ 1 \le s \le l_j$ and (57) holds with the sequential index ν removed where we define $\lambda_{js} = \lambda_j$ since $\lambda_{js}^\nu \xrightarrow{\nu} \lambda_j$ for $1 \le j \le m$ and $1 \le s \le l_j$.

Consequently, the same must be true for all of the sequences described in (52)–(58) where we denote their limits by removing the sequential index ν. Moreover, due to the continuity of f' and f'', all of these limits satisfy (52)–(58) with the sequential index ν removed.

Set

$$\mathcal{I} = \{(j, s) \mid \gamma_{js} > 0\} \quad \text{and} \quad \hat{\mathcal{I}} = \{j \mid (j, s) \in \mathcal{I}\}, \tag{59}$$

and note that $\hat{\mathcal{I}} \subset \mathcal{I}(p)$. Set $\gamma_j = \sum_{s=1}^{n_{js}} \gamma_{js},\ 1 \le j \le m$, so that

$$\sum_{j=1}^{m} \gamma_j = 1 \quad \text{with } \gamma_j > 0 \text{ for } j \in \hat{\mathcal{I}} \tag{60}$$

and $\gamma_j = 0$ otherwise. Write

$$w_j = \sum_{s=1}^{n_j} \omega_{js} e_{(n_j - s, \lambda_j)}, \quad 1 \le j \le m.$$

Since $u = \Psi^* w$, we have $\psi^*_{(\pi_j, \lambda_j)} w_j = (u_{j1}, \ldots, u_{jl_j}),\ 1 \le j \le m$, or equivalently,

$$\sum_{t=1}^{n_{js}} \omega_{jt} e_{(n_{jt} - t, \lambda_j)} = \sum_{t=1}^{n_{js}} \gamma_{js} \mu_{jst} e_{(n_{js} - t, \lambda_j)}, \quad 1 \le j \le m,\ 1 \le s \le l_j.$$

Therefore,

$$\omega_{jt} = \gamma_{js} \mu_{jst}, \quad 1 \le j \le m,\ 1 \le s \le l_j,\ 1 \le t \le n_{js}. \tag{61}$$

For $t = 1$, the first condition in (58) and the fact that $\gamma_{js} = 0$ for $(j, s) \notin \mathcal{I}$, gives

$$\omega_{j1} = \gamma_{js} \mu_{js1} = \frac{\gamma_{js}}{n_{js}} f'(\lambda_j) \quad 1 \le j \le m,\ 1 \le s \le l_j. \tag{62}$$

Since $f'(\lambda_j) \ne 0$ for $j \in \mathcal{I}(p)$, this gives $\gamma_{js}/n_{js} = \tau_{j1}$ for some $\tau_{j1} \in \mathbb{C}$, $j \in \mathcal{I}(p)$, $1 \le s \le l_j$. Therefore, $\gamma_{js} = \tau_{j1} n_{js},\ j \in \mathcal{I}(p),\ 1 \le s \le l_j$. Summing over s gives

$$\gamma_j = \sum_{s=1}^{l_j} \gamma_{js} = \sum_{s=1}^{l_j} \tau_{j1} n_{js} = \tau_{j1} n_j, \quad j \in \mathcal{I}(p),$$

that is, $\tau_{j1} = \frac{\gamma_j}{n_j},\ j \in \mathcal{I}(p)$. Hence, (62) implies that

$$\omega_{j1} = \frac{\gamma_j}{n_j} f'(\lambda_j), \quad 1 \le j \le m, \tag{63}$$

since $\gamma_{js} = 0 = \gamma_j$ for $j \notin \mathcal{I} \subset \mathcal{I}(p)$.

For $t = 2$, (61) tells us that $\omega_{j2} = \gamma_{js} \mu_{js2},\ 1 \le j \le m,\ 1 \le s \le l_j$. Multiplying each of these expressions by n_{js} and then summing over s gives

$$n_j \omega_{j2} = \sum_{s=1}^{l_j} n_{js} \omega_{j2} = \sum_{s=1}^{l_j} n_{js} \gamma_{js} \mu_{js2}, \quad 1 \le j \le m.$$

Combining this with the second condition in (58) for each $j \in \hat{\mathcal{I}}$, where $\hat{\mathcal{I}}$ is defined in (59), and using the definition of the sets $\mathcal{K}(n_{js}, \lambda_j)$ in (44), gives

$$\begin{aligned} n_j \left\langle \omega_{j2}, f'(\lambda_j)^2 \right\rangle &= \sum_{s=1}^{l_j} n_{js} \gamma_{js} \left\langle \mu_{js2}, f'(\lambda_j)^2 \right\rangle \\ &\le \sum_{s=1}^{l_j} \gamma_{js} \left\langle \mathrm{i} f'(\lambda_j), f''(\lambda_j)(\mathrm{i} f'(\lambda_j)) \right\rangle \\ &= \gamma_j \left\langle \mathrm{i} f'(\lambda_j), f''(\lambda_j)(\mathrm{i} f'(\lambda_j)) \right\rangle \quad \text{for } 1 \le j \le m. \end{aligned}$$

Setting $\hat{\omega}_{j2} = \gamma_j^{-1}\omega_{j2}$ for $j \in \hat{\mathcal{I}}$ gives $\left\langle \hat{\omega}_{j2},\, f'(\lambda_j)^2 \right\rangle \le \left\langle \mathrm{i}f'(\lambda_j),\, f''(\lambda_j)(\mathrm{i}f'(\lambda_j)) \right\rangle / n_j$, or equivalently,

$$\hat{\omega}_{j2} \in \mathcal{K}(n_j, \lambda_j) \quad \text{with } \omega_{j2} = \gamma_j \hat{\omega}_{j2} \text{ for } 1 \le j \le m, \tag{64}$$

since $\gamma_j = 0 = \gamma_{js}$ for $(j, s) \notin \mathcal{I}$ so that $\omega_{j2} = 0$ for $j \notin \hat{\mathcal{I}}$ by (61). Therefore, (60), (63) and (64) combine to imply that $w \in D(p)$ giving $z \in \hat{\partial}\mathbf{f}(p)$. $\square$

7. The radius mapping for polynomials

Formulas for the subdifferential of the abscissa mapping $\mathbf{a}$ as well as its subdifferential regularity are established in [2]. In this section we provide similar results for the radius mapping $\mathbf{r}$, defined in the introduction, using the results of the previous sections as well as the techniques and subdifferential formulas for $\hat{\partial}\mathbf{r}(e_{(n,\lambda_0)})$ established in [9].

The modulus function $r(\zeta) = |\zeta|$ is convex on $\mathbb{C}$ and twice continuously differentiable on $\mathbb{C} \setminus \{0\}$, but r'' is not positive definite on $\mathbb{C}\setminus\{0\}$. Therefore, the results of the previous sections do not directly apply except at the origin. In [9] this problem is overcome by introducing the quadratic function

$$\mathbf{r}_2(p) = \frac{1}{2}\mathbf{r}(p)^2 = \max\left\{ \frac{1}{2}|\lambda|^2 \,\middle|\, \lambda \in \mathcal{R}(p) \right\},$$

and establishing a relationship between the variational properties of $\mathbf{r}_2$ and those of $\mathbf{r}$.

Lemma 7.1 ([9, Lemma 7]). *Let $p \in \mathcal{P}^n$ be any polynomial for which $\mathbf{r}(p) > 0$. Then*

$$T_{\mathrm{epi}(\mathbf{r})}(p, \mu) = \left\{ \left(v, \frac{\eta}{\mu} \right) \mid (v, \eta) \in T_{\mathrm{epi}(\mathbf{r}_2)}\left(p, \frac{1}{2}\mu^2 \right) \right\} \quad \text{and}$$
$$\widehat{N}_{\mathrm{epi}(\mathbf{r})}(p, \mu) = \left\{ (w, \mu\tau) \mid (w, \tau) \in \widehat{N}_{\mathrm{epi}(\mathbf{r}_2)}\left(p, \frac{1}{2}\mu^2 \right) \right\}.$$

This lemma enables the following characterization of the regular subdifferential in the one root case.

Theorem 7.2 ([9, Theorem 12]). *Given $\lambda \in \mathbb{C}$ set $p = e_{(n,\lambda)}$ and define*

$$\mathcal{K}_{\mathbf{r}_2}(n, \lambda) = \{\theta \mid \langle \theta,\, \lambda^2 \rangle \le |\lambda|^2/n\},$$
$$\Delta_{\mathbf{r}_2}(n, \lambda) = 0 \times \{-\lambda/n\} \times \mathcal{K}_{\mathbf{r}_2}(n, \lambda) \times \mathbb{C}^{n-2}, \quad \text{and}$$
$$\Delta_{\mathbf{r}}(n, \lambda) = \begin{cases} 0 \times \left(\frac{1}{n}\mathbb{B} \right) \times \mathbb{C}^{n-1} & \text{if } \lambda = 0, \\ \frac{1}{|\lambda|}\Delta_{\mathbf{r}_2}(n, \lambda) & \text{otherwise,} \end{cases}$$

where $\mathbb{B}$ is the closed unit ball in $\mathbb{C}$. Then $\hat{\partial}\mathbf{r}(p) = \tau^*_{(n,\lambda)}\Delta_{\mathbf{r}}(n, \lambda)$ and, if $\lambda \ne 0$, then

$$\tau^*_{(n,\lambda)}\Delta_{\mathbf{r}_2}(n, \lambda) = \hat{\partial}\mathbf{r}_2(p) = \mathbf{r}(p)\hat{\partial}\mathbf{r}(e_{(n,\lambda)}).$$

Moreover, $\mathrm{d}\mathbf{r}(e_{(n,\lambda)}) = \sigma_{\hat{\partial}\mathbf{r}(e_{(n,\lambda)})}$, *that is, given $w = \sum_{s=0}^{n} \omega_s e_{(n-s,\lambda)}$,*

$$\mathrm{d}\mathbf{r}(e_{(n,\lambda)})(w) = \begin{cases} \frac{1}{n}|\omega_1| & \text{if } \lambda = 0, \\ \frac{1}{n|\lambda|}\left[|\omega_2| - \langle \lambda,\, \omega_1 \rangle \right] & \text{otherwise,} \end{cases}$$

whenever $\left\langle \lambda,\, \sqrt{-\omega_2} \right\rangle = 0$ and $\omega_s = 0,\ s = 3, \ldots, n$, with $\mathrm{d}\mathbf{r}(e_{(n,\lambda)})(w) = +\infty$ otherwise.

By combining Theorems 5.3 and 7.2 with Lemma 7.1 one can derive representations for all of the variational objects studied in the previous sections for the radius mapping $\mathbf{r}$. We give one such result for the subdifferential. As in Theorem 5.3, we make use of the following sets: for $\mathbf{r}(p) > 0$ and $1 \le j \le m$, set

$$\mathcal{K}_{\mathbf{r}}(n_j, \lambda_j) = \{\theta \mid \langle \theta,\, \lambda_j^2 \rangle \le |\lambda_j|/n_j\}$$
$$\hat{\Delta}_{\mathbf{r}}(n_j, \lambda_j) = \left\{ -\frac{1}{n_j}\frac{\lambda_j}{|\lambda_j|} \right\} \times \mathcal{K}_{\mathbf{r}}(n_j, \lambda_j) \times \mathbb{C}^{n_j-2}$$
$$D^j_{\mathbf{r}}(p) = \mathcal{T}_p^*(0, \ldots, 0, \hat{\Delta}_{\mathbf{r}}(n_j, \lambda_j), 0, \ldots, 0)$$
$$D(p) = \mathrm{conv} \cup_{j \in \mathcal{I}(p)} D^j_{\mathbf{r}}$$

and, for $\mathbf{r}(p) = 0$,

$$D(p) = \{0\} \times \frac{1}{n}\mathbb{B} \times \mathbb{C}^{n-1}.$$

Theorem 7.3. *The radius mapping $\mathbf{r}$ is subdifferentially regular on $\mathcal{M}^n$. Moreover, for $\kappa \in \mathbb{C}\setminus\{0\}$ and $p \in \mathcal{M}_1^n$ the subdifferential of $q = \kappa p$ is given by $\partial\mathbf{r}(q) = \bar{\kappa}^{-1}\partial\mathbf{r}(p)$ where $\partial\mathbf{r}(p) = \{z \mid \nabla F_p(0)^* D(p)\}$. In particular, if the inner product on $\mathcal{P}^n$ is given by $\langle\cdot,\cdot\rangle_{(\mathcal{P}^n,p)}$ defined in (42), then*

$$\partial\mathbf{r}(p) = \left\{ z \;\middle|\; \begin{array}{l} z = \displaystyle\sum_{j=1}^{m} r_j \sum_{s=1}^{n_j} \mu_{js} e_{(n_j - s, \lambda_j)}, \quad \textit{where } \mu_{js} = 0 \;\forall j \notin \mathcal{I}(p), \\ \exists \{\gamma_j\}_{j\in\mathcal{I}(p)} \subset [0,1] \quad \textit{with} \displaystyle\sum_{j\in\mathcal{I}(p)} \gamma_j = 1 \\ \textit{such that } \mu_{j1} = -\dfrac{\gamma_j}{n_j}\dfrac{\lambda_j}{|\lambda_j|} \quad \textit{and} \quad \mathrm{Re}\left(\overline{\lambda_j^2}\mu_{j2}\right) \le \dfrac{\gamma_j}{n_j}|\lambda_j| \;\forall j \in \mathcal{I}(p) \end{array} \right\}$$

when $\mathbf{r}(p) > 0$; otherwise, $\partial\mathbf{r}(p) = \{\frac{\mu}{n} e_{(n-1,0)} + q \mid |\mu| \le 1,\ q \in \mathcal{P}^{n-2}\}$.

Proof. By Theorem 6.2, $\mathbf{r}_2$ is subdifferentially regular on $\mathcal{M}^n \setminus \{e_{(n,0)}\}$ and Theorem 5.3 provides the formula (49) for $\partial\mathbf{r}_2(p)$. The representation for the regular normal cone for $\mathrm{epi}(\mathbf{r}_2)$ given in Lemma 7.1 combined with the relation (3) implies that $\mathbf{r}$ inherits the subdifferential regularity of $\mathbf{r}_2$ on $\mathcal{M}^n \setminus \{e_{(n,0)}\}$ with its subdifferential given by the formulas as stated via Theorem 5.3. At the polynomial $p = e_{(n,0)}$ the expression for $\partial\mathbf{r}(e_{(n,0)})$ given above is precisely the expression for $\hat{\partial}\mathbf{r}(e_{(n,0)})$ as given by Theorem 5.3 since $\partial r(0) = \mathbb{B}$. Therefore, it only remains to establish the subdifferential regularity of $\mathbf{r}$ at $e_{(n,0)}$. For this we make use of a limiting argument similar to the one given in the proof of Theorem 6.2.

Let $p^\nu \xrightarrow{\nu} e_{(n,0)}$ and $z^\nu \xrightarrow{\nu} z$ with $z^\nu \in \hat{\partial}\mathbf{r}(p^\nu)$ for all $\nu = 1, 2, \ldots$. We need to show that $z \in \hat{\partial}\mathbf{r}(e_{(n,0)})$. By Lemma 5.4, we can assume with no loss in generality that $\{p^\nu\} \subset \mathcal{M}_1^n$. If $p^\nu = e_{(n,0)}$ for some infinite subsequence $\nu \in J \subset \{1, 2, \ldots\}$, then we are done. Therefore, we can assume with no loss in generality that

$$\mathbf{r}(p^\nu) > 0 \quad \text{for all } \nu = 1, 2, \ldots. \tag{65}$$

Taking $p = e_{(n,0)}, f = r$, and $\mathbf{f} = \mathbf{r}$, we have that the entire development (52) through (58) holds true with $m = 1$ and $n_1 = n$. By (65), we have $f'(\lambda_{1s}^\nu) = r'(\lambda_{1s}^\nu) = \lambda_{1s}^\nu/|\lambda_{1s}^\nu|$ for all $(1,s) \in \mathcal{I}_\nu$ and $\nu = 1, 2, \ldots$. The compactness of $\mathbb{B}$ implies that we can assume with no loss in generality that there exist $\phi_s \in \mathbb{B}$ such that $\lambda_{1s}^\nu/|\lambda_{1s}^\nu| \to \phi_s \in \mathbb{B}$ for all $(1,s) \in \mathcal{I}_\nu$. Set $\phi_s = 0$ for $(1,s) \notin \mathcal{I}_\nu$. Since there are only finitely many possibilities for the index set $\mathcal{I}_\nu$, we can assume with no loss in generality that there is an index set $\tilde{\mathcal{I}}$ such that $\mathcal{I}_\nu = \tilde{\mathcal{I}}$ for all $\nu = 1, 2, \ldots$. Since $\tilde{\mathcal{I}}$ is fixed, the compactness of the set of possible γ_{1s}^ν's implies that we can also assume with no loss in generality that there exist γ_{1s} such that $\gamma_{1s}^\nu \to \gamma_{1s}$ for $1 \le s \le l_1$ with $\sum_{s=1}^{l_1} \gamma_{1s} = 1$. Therefore, all of the sequences defined in (52)–(58) have limits which we denote by removing the sequential index ν in all but (58) where we now have

$$\mu_{1s1}^\nu \xrightarrow{\nu} \mu_{1s1} = \frac{1}{n_{1s}}\phi_s, \qquad \mu_{1st}^\nu \xrightarrow{\nu} \mu_{1st} \in \mathbb{C}, \quad 2 \le t \le n_{1s},\ 1 \le s \le l_1, \tag{66}$$

and $\lambda_{1s} = 0$ for $1 \le s \le l_1$. As in the proof of Theorem 6.2, write $w_1 = \sum_{s=1}^{n_1} \omega_{1s} e_{(n_1 - s, 0)}$. Since $u = \Psi^* w$, we have $\psi_{(\pi_1,0)}^* w_1 = (u_{11}, \ldots, u_{1l_1})$, or equivalently,

$$\sum_{t=1}^{n_{1s}} \omega_{1t} e_{(n_{1s}-t,0)} = \sum_{t=1}^{n_{1s}} \gamma_{1s}\mu_{1st} e_{(n_{1s}-t,0)}, \quad 1 \le s \le l_1.$$

Therefore, $\omega_{1t} = \gamma_{1s}\mu_{1st},\ 1 \le s \le l_1,\ 1 \le t \le n_{1s}$. For $t = 1$, the first condition in (66) and the fact that $\gamma_{1s} = 0 = \phi_s$ for $(1,s) \notin \mathcal{I}$, gives $\omega_{11} = \gamma_{1s}\mu_{1s1} = \frac{\gamma_{1s}}{n_{1s}}\phi_s,\ 1 \le s \le l_j$. Multiplying this expression through by n_{1s} and summing over s gives $n\omega_{11} = n_1\omega_{11} = \sum_{s=1}^{l_1} \gamma_{1s}\phi_s$, or equivalently, $\omega_{11} = \frac{1}{n}\sum_{s=1}^{l_1} \gamma_{1s}\phi_s \in \frac{1}{n}\mathbb{B}$, since $\sum_{s=1}^{l_1}\gamma_{1s} = 1$ and $\mathbb{B}$ is convex. That is, $z \in \hat{\partial}\mathbf{r}(e_{(n,0)})$. $\square$

References

[1] J.V. Burke, A.S. Lewis, M.L. Overton, Optimizing matrix stability, Proceedings of the American Mathematical Society 129 (6) (2000) 1635–1642.
[2] J.V. Burke, M.L. Overton, Variational analysis of the abscissa mapping for polynomials, SIAM Journal on Control and Optimization 39 (6) (2001) 1651–1676.
[3] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM, Philadelphia, 1990, Republished.
[4] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation, vol. 1, Springer, New York, 2005.
[5] R.T. Rockafellar, R.J.B. Wets, Variational Analysis, Springer, 1998.
[6] L.V. Levantovskii, On singularities of the boundary of a stability region, Moscow University Mathematics Bulletin 35 (1980) 19–22.

[7] J.V. Burke, A.S. Lewis, M.L. Overton, Variational analysis of the abscissa mapping for polynomials via the Gauss-lucas theorem, Journal of Global Optimization 28 (2004) 259–268.
[8] M. Marden, Geometry of Polynomials, American Mathematical Society, 1966.
[9] J.V. Burke, A.S. Lewis, M.L. Overton, Variational analysis of functions of the roots of polynomials, Mathematical Programming (2005) 263–292.
[10] Julia Eaton, Variational properties of polynomial root functions and spectral functions, Ph.D. Thesis, Department of Mathematics, University of Washington, 2010.
[11] E. Steinitz, Bedingt konvergente reihen und konvexe systeme, I, II, III, Journal für Reine und Angewandte Mathematik 143 (1913) 128–175. 144, 146; 1–40; 1–52; 1914, 1916.
[12] A. Auslender, M. Teboulle, Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer, New York, 2000.
[13] J.-B. Hiriart-Urruty, C. Lemaréchal, Fundamentals of Convex Analysis, Springer, Berlin, 2001.