On the Subdifferentiability of Functions of a Matrix Spectrum
II: Subdifferential Formulas

James V. Burke* and Michael L. Overton†

Abstract

Variational properties for the spectral radius and spectral abscissa of an analytic matrix valued mapping $A: \mathbb{C}^s \mapsto \mathbb{C}^{n\times n}$ are considered. A notion of directional differentiability is introduced that allows us to exploit the perturbation results of Part I of this study. Lower bounds for the directional derivative are established which yield formulas for the directional derivative when a natural nondegeneracy condition is satisfied. These formulas are interpreted in the extreme cases where the eigenvalues attaining either the spectral radius or the spectral abscissa are nonderogatory or semisimple (nondefective). We conclude by investigating the relationship with the proximal normal subdifferential. In particular, it is shown that the proximal normal subdifferential is always nonempty.

1 Introduction

In this article, we develop a few of the variational properties of the spectral radius, ρ, and the spectral abscissa, α, for the analytic matrix valued mapping $A: \mathbb{C}^s \mapsto \mathbb{C}^{n\times n}$. The mappings ρ and α are given by the formulas

$$\rho(z) := \max\{|\lambda|: \lambda \in \Sigma(z)\},$$

and

$$\alpha(z) := \max\{\mathrm{Re}\,\lambda : \lambda \in \Sigma(z)\},$$

where the multifunction $\Sigma: \mathbb{C}^s \mapsto \mathbb{C}^{n\times n}$ is the spectrum of $A(z)$, i.e. the roots of the characteristic polynomial

$$P(z,\lambda) = \det(\lambda I - A(z)).$$

*Department of Mathematics, University of Washington. The work of this author was supported in part by National Science Foundation Grant DMS-88-03206

†Computer Science Department, Courant Institute of Mathematical Sciences, New York University. The work of this author was supported in part by National Science Foundation Grant CCR-88-02408.

As discussed in Part I [3], the mappings α and ρ, although continuous, are not, in general, lipschitzian near values of z for which P has multiple roots. The nonlipschitzian behavior arises from the bifurcation of these multiple roots. Along analytic curves in $\mathbb{C}^s$ these roots are algebroidal functions and as such can be described as the branches of an analytic function with each branch being representable as a power series in fractional powers of the perturbation. It is these classical results that inspire us to introduce a notion of directional differentiability which depends in an essential way on the underlying analyticity of A.

For $f: \mathbb{C}^s \mapsto \mathbb{R}$, define $f^h(z;\cdot): \mathbb{C}^s \mapsto \mathbb{R} \cup \{\pm\infty\}$ by

$$f^h(z;d) := \inf_{\gamma \in \Gamma(z,d)} \liminf_{\epsilon \downarrow 0} \frac{f(\gamma(\epsilon)) - f(z)}{\epsilon}, \tag{1}$$

where

$$\Gamma(z,d) := \{\gamma: \mathbb{C} \mapsto \mathbb{C}^s \ : \ \gamma \ \text{is analytic}\ , \ \gamma(0) = z, \ \text{and}\ \gamma'(0) = d\}. \tag{2}$$

The notation $\epsilon \downarrow 0$ implies that only *real* (and positive) values of ϵ are to be considered. The superscript h in (1) is used to emphasize that $f^h(z;d)$ depends only on the analytic (holomorphic) curves in $\mathbb{C}^s$ that pass through z. At first glance it may appear that this notion of directional differentiability is not sufficiently robust to capture the full range of first order variational behavior that is possible for nonlipschitzian functions. However, as we hope to demonstrate, it *is* sufficiently robust for the analysis of the functions ρ and α. In order to see why this is so, we need to recall certain properties of the Puiseux-Newton series expansions for the roots of P, which were introduced by way of example in the discussion following Lemma 1 in Part I of this study. Here we provide a little more detail.

Let z, d be in $\mathbb{C}^s$ and choose $\gamma \in \Gamma(z,d)$. The roots of $P(\gamma(\epsilon),\lambda)$ are given by series expansions in fractional powers of ϵ, that is, Puiseux-Newton series of the form

$$\lambda(\epsilon) := \lambda_0 + \kappa\epsilon^p + \cdots, \tag{3}$$

where $p := q/r \geq 1/n$ for some pair of relatively prime integers q and r. Associated with each such value of p there are r series of the form (3) yielding roots of $P(\gamma(\epsilon),\lambda)$ with κ taking the values

$$\eta^{1/r}\omega^j \quad \text{for } j = 1, \cdots, r,$$

for some $\eta \in \mathbb{C}$ where $\eta^{1/r}$ is the principal rth root of η and ω is the principal rth root of unity. An immediate consequence of this type of *splitting* of the roots of P is that, when evaluating the directional derivative (1) for the spectral abscissa and radius, one can replace the limit infimum in the definition with a limit. Moreover, again in the case of ρ and α, this directional derivative always exists, possibly taking the value $+\infty$ but never the value $-\infty$. This structure also indicates why the spectral radius and abscissa are Hölder continuous of order at least $1/n$. Thus, a natural extension to (1) would be to consider *Hölder* type directional derivatives based on fractional powers of ϵ. Although it is possible to pursue such an approach using the techniques developed in Part 1 of this study, we do not do so here.

2 The Directional Derivative

In this section, we apply Lemmas 4 and 5 of Part I of this study to obtain information about the directional derivative defined in (1) for the spectral abscissa and radius. We continue to assume that $A\colon \mathbb{C}^s \mapsto \mathbb{C}^{n\times n}$ is an analytic matrix function and $z_0 \in \mathbb{C}^s$. Also let $A(z_0)$ have the block diagonal decomposition

$$A(z_0) = A^{(0)} = PDP^{-1}, \quad D = \mathrm{Diag}(D_1, \ldots, D_m), \tag{4}$$

where m is the number of distinct eigenvalues of $A(z_0)$ and the kth diagonal block D_k is upper triangular with constant diagonal λ_k, i.e.

$$D_k = \lambda_k I + N_k,$$

with N_k strictly upper triangular and hence nilpotent. Let n_k denote the order of D_k, i.e., the multiplicity of the eigenvalue λ_k. Suppose

$$\frac{\partial A}{\partial z_l}(z_0) = A_l^{(1)}, \text{ for } l = 1, \cdots, s,$$

and let $B_{1l}, \ldots, B_{ml}$ be the corresponding diagonal blocks of $P^{-1}A_l^{(1)}P$ for $l = 1, \cdots, s$. With this notation established, we have the following results.

Theorem 1 *Choose $d \in \mathbb{C}^s$. If for some $k \in \mathcal{A}(z_0) := \{k : \lambda_k \in \Sigma(z_0),\ \mathrm{Re}\,\lambda_k = \alpha(z_0)\}$ any one of the conditions*

$$\mathrm{Re}\sum_{l=1}^{s}(\mathrm{tr}\ N_k B_{kl})d_l \le 0,\ \mathrm{Im}\sum_{l=1}^{s}(\mathrm{tr}\ N_k B_{kl})d_l = 0, \tag{5}$$

$$\sum_{l=1}^{s}(\mathrm{tr}\ N_k^{j-1} B_{kl})d_l = 0,\ j = 3, \ldots, n_k, \tag{6}$$

is violated, then

$$\alpha^h(z_0; d) = +\infty\ ; \tag{7}$$

otherwise,

$$\alpha^h(z_0; d) \ge \phi(z_0, d), \tag{8}$$

where

$$\phi(z_0, d) := \max\left\{\frac{\mathrm{Re}\ \sum_{l=1}^{s}(\mathrm{tr}\ B_{kl})d_l}{n_k}\ :\ k \in \mathcal{A}(z_0)\right\}. \tag{9}$$

Moreover, if conditions (5) and (6) hold and the vectors

$$\{[\mathrm{tr}\ N_k^{j-1}B_{k1}, \cdots,\ \mathrm{tr}\ N_k^{j-1}B_{ks}]^T :\ j = 1, \ldots, n_k,\ k \in \mathcal{A}_1(z_0, d)\}$$

are linearly independent, where $\mathcal{A}_1(z_0, d)$ is the set of indices in $\mathcal{A}(z_0)$ which achieve the maximum in (9), then equality holds in (8).

Partial Proof We will establish the first statement by applying Lemma 4 of Part I, using the correspondence between $A^{(1)}$ in Part I and $\sum_{l=1}^{s} d_l A_l^{(1)}$ here. It follows immediately that (7) holds if (5) or (6) is violated. Indeed, if for some $k \in \mathcal{A}(z_0)$ one of the conditions (5) or (6) fail to hold, then for every $\gamma \in \Gamma(z_0, d)$ the inequality

$$\text{Re }(\lambda(\epsilon) - \lambda_k) \leq \delta\epsilon + o(\epsilon)$$

fails to hold for every $\delta \in \mathcal{R}$, for some branch $\lambda(\epsilon)$ of the splitting of λ_k along the curve γ. Consequently, the limit defining $\alpha^h(z_0; d)$ is $+\infty$ for every $\gamma \in \Gamma(z_0, d)$ and so $\alpha^h(z_0; d) = +\infty$. On the other hand, if $\delta > \alpha^h(z_0; d)$, then there is a $\gamma \in \Gamma(z_0, d)$ such that

$$\lim_{\epsilon \downarrow 0} \frac{\alpha(\gamma(\epsilon)) - \alpha(z_0)}{\epsilon} < \delta,$$

or equivalently,

$$\text{Re }(\lambda(\epsilon) - \lambda_k) < \delta\epsilon \text{ for } \epsilon \in [0, \epsilon_0],$$

for some $\epsilon_0 > 0$, for every branch $\lambda(\epsilon)$ of the splitting of λ_k along the curve γ ,and for every $k \in \mathcal{A}(z_0)$. Now by applying Lemma 4 of Part I to this inequality for each $k \in \mathcal{A}(z_0)$ and then letting $\delta \downarrow \alpha^h(z_0; d)$, we obtain inequality (8). The proof of the last statement requires the implicit function theorem and is given in [4]. □

Theorem 2 *Choose $d \in \mathbb{C}^s$ and set*

$$\mathcal{R}(z_0) := \{k : |\lambda_k| = \rho(z_0)\}.$$

We will consider two cases; $\rho(z_0) = 0$ and $\rho(z_0) \neq 0$.

1) Let us assume that $\rho(z_0) = 0$ so that $m = 1$ and $\lambda_1 = 0$. If any one of the conditions

$$\sum_{l=1}^{s} (\text{tr } N_1^{j-1} B_{1l}) d_l = 0, \; j = 2, \ldots, n_k, \tag{10}$$

is violated, then

$$\rho^h(z_0; d) = +\infty \;;$$

otherwise,

$$\rho^h(z_0; d) \geq \xi(z_0, d), \tag{11}$$

where

$$\xi(z_0, d) := \frac{1}{n} |\sum_{l=1}^{s} (\text{tr } B_{1l}) d_l| \;.$$

Moreover, if condition (10) holds and the vectors

$$\{[\text{tr } N_1^{j-1} B_{11}, \cdots, \text{ tr } N_1^{j-1} B_{1s}]^T : \; j = 1, \ldots, n\}$$

are linearly independent, then equality holds in (11).

2) Let us assume that $\rho(z_0) \neq 0$. If for some $k \in \mathcal{R}(z_0)$ any one of the conditions

$$\text{Re } \overline{\lambda}_k^{\,2} \sum_{l=1}^{s} (\text{tr } N_k B_{kl}) d_l \leq 0, \ \text{Im } \overline{\lambda}_k^{\,2} \sum_{l=1}^{s} (\text{tr } N_k B_{kl}) d_l = 0, \tag{12}$$

$$\sum_{l=1}^{s} (\text{tr } N_k^{j-1} B_{kl}) d_l = 0, \ j = 3, \ldots, n_k, \tag{13}$$

is violated, then

$$\rho^h(z_0; d) = +\infty \ ;$$

otherwise,

$$\rho^h(z_0; d) \geq \psi(z_0, d), \tag{14}$$

where

$$\psi(z_0, d) := \max\{\frac{1}{n_k \rho(z_0)} [\text{Re } (\overline{\lambda}_k \sum_{l=1}^{s} (\text{tr } B_{kl}) d_l) + |\sum_{l=1}^{s} (\text{tr } N_k B_{kl}) d_l|] \ : \ k \in \mathcal{R}(z_0)\}. \tag{15}$$

Moreover, if conditions (12) and (13) hold and the vectors

$$\{[\text{tr } N_k^{j-1} B_{k1}, \cdots, \ \text{tr } N_k^{j-1} B_{ks}]^T \ : \ j = 1, \ldots, n_k, \ k \in \mathcal{R}_1(z_0, d)\}$$

are linearly independent, where $\mathcal{R}_1(z_0, d)$ is the set of indices in $\mathcal{R}(z_0)$ achieving the maximum in (15), then equality holds in (14).

Partial Proof First consider $\rho(z_0) = 0$. It follows immediately from Lemma 5 of Part I that (10) must hold for $j = 2$ as well as $j = 3, \cdots, n$, if the directional derivative is to be finite valued. Now assume that (10) holds and suppose $\delta \geq \rho^h(z_0; d)$. Then there is a $\gamma \in \Gamma(z_0, d)$ such that

$$|\sum \lambda(\epsilon)| \leq \sum |\lambda(\epsilon)| < n\delta\epsilon \ ,$$

where the sum is taken over all branches $\lambda(\epsilon)$. Letting $\delta \downarrow 0$ yields (11) in view of Lemma 3 in Part I.

Next assume that $\rho(z_0) \neq 0$. The proof of inequality (14) is a straight forward application of Lemma 5 in Part I. Observe that if for some $k \in \mathcal{R}(z_0)$ one of the conditions (12) or (13) fail to hold, then for every $\gamma \in \Gamma(z_0, d)$ the inequality

$$|\lambda(\epsilon)| - |\lambda_k| \leq \delta\epsilon + o(\epsilon)$$

fails to hold for every $\delta \in \mathcal{R}$, for some branch $\lambda(\epsilon)$ of the splitting of λ_k along the curve γ. Consequently, the limit defining $\rho^h(z_0; d)$ is $+\infty$ for every $\gamma \in \Gamma(z_0, d)$ and so $\rho^h(z_0; d) = +\infty$. On the other hand, if $\delta > \rho^h(z_0; d)$, then there is a $\gamma \in \Gamma(z_0, d)$ such that

$$\lim_{\epsilon \downarrow 0} \frac{\rho(\gamma(\epsilon)) - \rho(z_0)}{\epsilon} < \delta,$$

or equivalently,

$$|\lambda(\epsilon)| - |\lambda_k| < \delta\epsilon \text{ for } \epsilon \in [0, \epsilon_0],$$

for some $\epsilon_0 > 0$, for every branch $\lambda(\epsilon)$ of the splitting of λ_k along the curve γ ,and for every $k \in \mathcal{R}(z_0)$. Now by applying Lemma 5 of Part I to this inequality for each $k \in \mathcal{A}(z_0)$ and then letting $\delta \downarrow \rho^h(z_0; d)$, we obtain inequality (14).

The proofs of equality in both cases makes use of the implicit function theorem and are given in [4]. □

3 The Nonderogatory and Semisimple Cases

In this section we consider two cases that illustrate the results in Theorems 1 and 2. Let $\lambda_k \in \Sigma(z_0)$ have multiplicity n_k. Then λ_k is said to be *nonderogatory* if it possesses only a single eigenvector. The nonderogatory multiple eigenvalues are, in a quantifiable sense, the most generic kind of multiple eigenvalue [1]. The eigenvalue λ_k is said to be *semisimple* if it is associated with n_k independent eigenvectors.

Now suppose $A(z_0)$ has block diagonal form as described in the previous section. Then λ_k is semisimple if and only if $N_k = 0$, and λ_k is nonderogatory if and only if $N_k^{n_k-1} \neq 0$. Therefore, if λ_k is derogatory, then for some $j_0 < n_k$ we have that $N_k^j = 0$ for $j \geq j_0$. This observation is significant in the interpretation of Theorems 1 and 2. These theorems provide formulas for the directional derivatives $\alpha^h(z_0; d)$ and $\rho^h(z_0; d)$ when it is known that vectors given by

$$[\text{tr } N_k^{j-1} B_{k1}, \cdots, \text{ tr } N_k^{j-1} B_{ks}]^T, \text{ for } j = 1, \ldots, n_k$$

are linearly independent. However these vectors cannot be linearly independent if the eigenvalue λ_k is derogatory since in this case $N_k^{j-1} = 0$ for at least one value of j satisfying $1 \leq j \leq n_k$. Thus the formulas that are provided in Theorems 1 and 2 are at best sharp only in the nonderogatory case. Such an outcome could have been predicted from the outset. As has been indicated, our results are derived from results for polynomials and are extended to matrices by application to the characteristic polynomial. Such results are most applicable to nonderogatory eigenvalues since the multiplicity of these eigenvalues in the characteristic polynomial agrees with their multiplicity in the minimal polynomial. In this case, the characteristic polynomial accurately reflects the behavior of the eigenvalue. This is not true for derogatory eigenvalues.

It is instructive to see just how far off these lower bounds can be by considering the semisimple case. Semisimple and nonderogatory eigenvalues lie at opposite structural extremes whenever $n_k > 1$ (of course, if $n_k = 1$, then the eigenvalue is both semisimple and nonderogatory). Thus, we would expect the lower bounds for the directional derivatives $\alpha^h(z_0; d)$ and $\rho^h(z_0; d)$ to be most inaccurate in this case. By employing the resolvent theory for linear operators [5] to the semisimple case, one can obtain precise formulas for the required directional derivatives. This is done in [7, Lemma 3.5]. As an immediate consequence of this result we obtain the following theorem.

Theorem 3 *Suppose $\mathcal{A}(z_0)$ and $\mathcal{R}(z_0)$ contain only semisimple eigenvalues of $A(z_0)$. Let*

$d \in \mathbb{C}^s$ and denote by λ'_{kl}, for $l = 1, \ldots, n_k$, the eigenvalues of $\sum_{l=1}^s B_{kl} d_l$ for each $k \in \mathcal{A}(z_0)$.

1) The spectral abscissa: The ordinary directional derivative $\alpha'(z_0; d)$ exists and satisfies

$$\alpha'(z_0; d) = \alpha^h(z_0; d) = \max_{k \in \mathcal{A}(z_0)} \max_{1 \leq l \leq n_k} \{\mathrm{Re}\, \lambda'_{kl}\}. \tag{16}$$

2) The spectral radius: The ordinary directional derivative $\rho'(z_0; d)$ exists and satisfies

$$\rho'(z^0; d) = \rho^h(z_0; d) = \begin{cases} \max_{k \in \mathcal{R}(z_0)} \max_{1 \leq l \leq n_k} |\lambda'_{kl}| & \text{if } \rho(z_0) = 0, \\ \frac{1}{\rho(z_0)} \max_{k \in \mathcal{R}(z_0)} \max_{1 \leq l \leq n_k} \mathrm{Re}\, \overline{\lambda}_k \lambda'_{kl} & \text{if } \rho(z_0) \neq 0. \end{cases} \tag{17}$$

Let us suppose that both $\mathcal{A}(z_0)$ and $\mathcal{R}(z_0)$ contain the single element k and that the associated eigenvalue λ_k is semisimple. Regarding the spectral abscissa, formula (16) says that $\alpha^h(z_0; d)$ is the maximum of the real parts of the eigenvalues of the matrix

$$\sum_{l=1}^{s} B_{kl} d_l \tag{18}$$

whereas the lower bound for $\alpha^h(z_0; d)$ given by the right hand side of inequality (8) in Theorem 1 is the average of these quantities. This behavior is repeated in the case of the spectral radius. If $\rho(z_0) = 0$, then $\rho^h(z_0; d)$ is the spectral radius of the matrix (18) while the lower bound given by (11) is the modulus of the average of the eigenvalues of (18). Similarly, if $\rho(z_0) \neq 0$, then $\rho^h(z_0; d)$ is the maximum of the values

$$\frac{1}{\rho(z_0)} \mathrm{Re}\, \overline{\lambda}_k \lambda'_{kl}$$

while the lower bound in (14) is the average of these values (since $N_k = 0$). Apparently the difference between the actual value of the directional derivative (1) and the lower bounds given by Theorems 1 and 2 can be quite great. Nonetheless, these lower bounds indicate when these directional derivatives are infinite-valued and in the generic nonderogatory case they give precise values for the directional derivative.

In the general case, when the eigenvalues are derogatory but not semisimple, little precise information is known beyond the general lower bounds provided by Theorems 1 and 2. The semisimple case illustrated above indicates a possible upper bound.

4 The Proximal Normal Subdifferential

In this section, we use our knowledge of the directional derivatives $\alpha^h(z; d)$ and $\rho^h(z; d)$ to obtain insight into the structure of the proximal normal subdifferential for the spectral radius and spectral abscissa. A vector $v \in \mathbb{C}^s$ is said to be a proximal normal subgradient for the function $f\colon \mathbb{C}^s \mapsto \mathcal{R}$ at $z_0 \in \mathbb{C}^s$ if

$$\liminf_{z \to z_0} \frac{f(z) - f(z_0) - \langle v, z - z_0 \rangle}{\|z - z_0\|^2} > -\infty, \tag{19}$$

where the inner product on $\mathbb{C}^s$ is given by

$$\langle y, w \rangle := \mathrm{Re}\ \overline{y}^T w.$$

The collection of all proximal normal subgradients at z_0 is called the proximal normal subdifferential of f at z_0 and is denoted by $\partial_2 f(z_0)$. Our interest in understanding the proximal normal subdifferential for the functions α and ρ stems from the rich calculus that is available for this subdifferential and its limits. For more information about this subdifferential, we refer the reader to Mordukhovich [6] and Rockafellar [8].

Let $z_0 \in \mathbb{C}^s$ and suppose that $A(z_0)$ has the block diagonal decomposition described in Section 2. According to Theorem 1, the directional derivative $\alpha^h(z_0; d)$ is bounded below by the support function for the set

$$\begin{aligned} \Phi(z_0) \quad := \quad & \{ \frac{1}{n_k} [\mathrm{tr}\ \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{B}_{ks}]^T \ :\ k \in \mathcal{A}(z_0) \ \} \\ + \quad & [\{ \zeta [\mathrm{tr}\ \overline{N}_k \overline{B}_{k1}, \cdots, \mathrm{tr}\ \overline{N}_k \overline{B}_{ks}]^T \ :\ k \in \mathcal{A}(z_0),\ \mathrm{Re}\ \zeta \geq 0 \ \} \cup \\ & \mathrm{span}\{ [\mathrm{tr}\ \overline{N}_k^{j-1} \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{N}_k^{j-1} \overline{B}_{ks}]^T \ :\ j = 3, \ldots, n_k,\ k \in \mathcal{A}(z_0) \ \}], \end{aligned}$$

where the support functional for a set $\Omega \subset \mathbb{C}^s$ is defined by the expression

$$\sigma_\Omega(d) := \sup_{y \in \Omega} \langle y, d \rangle.$$

Similarly, by Theorem 2, the directional derivative $\rho^h(z_0; d)$ is bounded below by the support functional for the set

$$\begin{aligned} \Xi(z_0) \quad := \quad & \{ \frac{e^{i\theta}}{n} [\mathrm{tr}\ \overline{B}_{11}, \cdots,\ \mathrm{tr}\ \overline{B}_{1s}]^T \ :\ \theta \in [0, 2\pi] \ \} \\ + \quad & \mathrm{span}\{ [\mathrm{tr}\ \overline{N}_1^{j-1} \overline{B}_{11}, \cdots,\ \mathrm{tr}\ \overline{N}_1^{j-1} \overline{B}_{1s}]^T \ :\ j = 2, \ldots, n \}, \end{aligned}$$

when $\rho(z_0) = 0$, and is bounded below by the support functional for the set

$$\begin{aligned} \Psi(z_0) \quad := \quad & \left\{ \begin{array}{c|c} \frac{e^{i\theta}}{n_k \rho(z_0)} [\mathrm{tr}\ \overline{N}_k \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{N}_k \overline{B}_{ks}]^T & \\ & k \in \mathcal{R}(z_0, d),\ \theta \in [0, 2\pi] \\ + \frac{\lambda_k}{n_k \rho(z_0)} [\mathrm{tr}\ \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{B}_{ks}]^T & \end{array} \right\} \\ + \quad & [\{ \zeta \lambda_k^2 [\mathrm{tr}\ \overline{N}_k \overline{B}_{k1}, \cdots, \mathrm{tr}\ \overline{N}_k \overline{B}_{ks}] \ :\ k \in \mathcal{R}(z_0),\ \mathrm{Re}\ \zeta \geq 0 \ \} \cup \\ & \mathrm{span}\{ [\mathrm{tr}\ \overline{N}_k^{j-1} \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{N}_k^{j-1} \overline{B}_{ks}]^T \ :\ j = 3, \ldots, n_k,\ k \in \mathcal{R}(z_0) \ \}], \end{aligned}$$

when $\rho(z_0) \neq 0$. The vectors defining the sets Φ, Ξ, and Ψ are our prime candidates for proximal normal subgradients. The partial results that we have been able to obtain to date depend upon the resolvent theory for analytic operator valued mappings on $\mathbb{C}^s$ developed by Baumgärtel [2]. The key result that we require in this regard is [2, Theorem 2, page 374]. In order to understand this result, one must first understand the notion of a λ_k-group for $\lambda_k \in \Sigma(z_0)$.

Given $\lambda_k \in \Sigma(z_0)$ the λ_k-group is the set of all functions $\lambda(z) \in \Sigma(z)$ such that $\lambda(z) \to \lambda_k$ as $z \to z_0$. The λ_k-groups are well defined since the spectrum $\Sigma(z)$ is a continuous multifunction. The projection onto the invariant subspace for the entire λ_k-group is called the group projection for λ_k. The following remarkable fact, established in [2, Theorem 2, page 374], states that each such group projection is locally analytic.

Theorem 4 *Let $z_0 \in \mathbb{C}^s$. Then there is a neighborhood Ω of z_0 in $\mathbb{C}^s$ such that for each $\lambda_k \in \Sigma(z_0)$ the group projection $\Pi_k(z)$ associated with λ_k is analytic on Ω. Moreover, for each $\lambda_k \in \Sigma(z_0)$ there is a neighborhood G_k of λ_k in $\mathbb{C}$ such that*

$$Spec(A(z)\Pi_k(z)) = \Sigma(z) \cap G_k$$

for all $z \in \Omega$ and $\lambda_k \in \Sigma(z_0)$ where $Spec(A(z)\Pi_k(z))$ denotes the spectrum of $A(z)\Pi_k(z)$.

This result is the key to establishing the local factorization of the characteristic polynomial used in Section 2 of Part I of this study. In the current context, we simply observe that Theorem 4 implies the analyticity of the trace of each λ_k-group,

$$\tau_k(z) := \text{tr } A(z)\Pi_k(z), \tag{20}$$

since this trace is simply minus the coefficient of the $(\lambda - \lambda_k)^{n_k - 1}$ term in the characteristic polynomial for $A(z)P_k(z)$. From Lemma 3 in Part I, we find that the gradient of the trace of each λ_k-group at z_0 is given by

$$\nabla \tau_k(z_0) = [\text{tr } B_{k1}, \cdots, \text{ tr } B_{ks}]^T. \tag{21}$$

By combining this fact with the observation that the maximum of a finite set of real numbers always exceeds (or equals) their average, we are able to establish that certain members of the sets $\Phi(z_0)$, $\Xi(z_0)$ (if $\rho(z_0) = 0$), and $\Psi(z_0)$ (if $\rho(z_0) \neq 0$) are proximal normal subgradients.

Theorem 5 *Let $z_0 \in \mathbb{C}^s$ and suppose that $A(z_0)$ has the block diagonal decomposition described in Section 2. Then*

$$\frac{1}{n_k}[\text{tr } \overline{B}_{k1}, \cdots, \text{ tr } \overline{B}_{ks}]^T \in \partial_2 \alpha(z_0), \tag{22}$$

for every $k \in \mathcal{A}(z_0)$.

Proof Let $k \in \mathcal{A}(z_0)$. Since $\tau_k(z)$ is analytic near z_0 and $\alpha(z) \geq (1/n_k)\tau_k(z)$ for z near z_0 and $k \in \mathcal{A}(z_0)$, we have that

$$\begin{aligned} \alpha(z) - \alpha(z_0) - \frac{1}{n_k}\text{Re } \nabla \tau_k(z_0)^T(z - z_0) &\geq \frac{1}{n_k}\text{Re } [\tau_k(z) - \tau_k(z_0) - \nabla \tau_k(z_0)^T(z - z_0)] \\ &= O(\|z - z_0\|^2). \end{aligned}$$

Thus, by (21)

$$\liminf_{z \to z_0} \frac{\alpha(z) - \alpha(z_0) - \langle \frac{1}{n_k}[\mathrm{tr}\ \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{B}_{ks}], z - z_0 \rangle}{\|z - z_0\|^2} > -\infty.$$

□

Theorem 6 *Let $z_0 \in \mathbb{C}^s$ and suppose that $A(z_0)$ has the block diagonal decomposition described in Section 2.*

1. *Suppose that $\rho(z_0) = 0$ so that $\Sigma(z_0)$ reduces to the single eigenvalue $\lambda_1 = 0$. Then*

$$\frac{\eta e^{i\theta}}{n}[\mathrm{tr}\ \overline{B}_{11}, \cdots,\ \mathrm{tr}\ \overline{B}_{1s}]^T \in \partial_2 \rho(z_0), \tag{23}$$

for every $\eta \in [0,1]$ and $\theta \in [0, 2\pi]$.

2. *Suppose that $\rho(z_0) \neq 0$. Then*

$$\frac{\lambda_k}{n_k \rho(z_0)}[\mathrm{tr}\ \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{B}_{ks}]^T \in \partial_2 \rho(z_0), \tag{24}$$

for each $k \in \mathcal{R}(z_0)$.

Proof (1) Let $\eta \in [0,1]$ and $\theta \in [0, 2\pi]$ be given. Since $\tau_1(z)$ is analytic near z_0, we have that

$$\begin{aligned} \rho(z) - \rho(z_0) &- \langle \frac{\eta e^{i\theta}}{n}[\mathrm{tr}\ \overline{B}_{k1}, \cdots,\ \mathrm{tr}\ \overline{B}_{ks}], z - z_0 \rangle \\ &\geq \frac{1}{n}[|\tau_1(z)| - |\tau_1(z_0) + \nabla \tau_1(z_0)^T (z - z_0)|] \\ &= O(\|z - z_0\|^2), \end{aligned}$$

which establishes the result.
(2) Let $k \in \mathcal{R}(z_0)$. Since $\tau_k(z)$ is analytic near z_0, we have that

$$\begin{aligned} \rho(z) - \rho(z_0) &\geq \frac{1}{n_k}[|\tau_k(z)| - |\tau_k(z_0)|] \qquad (25) \\ &= \frac{|\tau_k(z)|^2 - |\tau_k(z_0)|^2}{n_k[|\tau_k(z)| + |\tau_k(z_0)|]} \\ &= \frac{2\mathrm{Re}\ (\overline{\tau}_k(z_0)(\tau_k(z) - \tau_k(z_0))) + |\tau_k(z) - \tau_k(z_0)|^2}{n_k[|\tau_k(z)| + |\tau_k(z_0)|]} \\ &\geq \frac{2\mathrm{Re}\ (\overline{\tau}_k(z_0)(\tau_k(z) - \tau_k(z_0)))}{n_k[|\tau_k(z)| + |\tau_k(z_0)|]} \\ &= \frac{2\mathrm{Re}\ (\overline{\tau}_k(z_0)(\nabla \tau_k(z_0)^T (z - z_0) + O(\|z - z_0\|^2)))}{n_k[|\tau_k(z)| + |\tau_k(z_0)|]} \\ &= \frac{2\mathrm{Re}\ \overline{\tau}_k(z_0)(\nabla \tau_k(z_0)^T (z - z_0))}{2n_k\ |\tau_k(z_0)|} + O(\|z - z_0\|^2) \qquad (26) \\ &= \frac{\mathrm{Re}\ \overline{\lambda}_k \nabla \tau_k(z_0)(z - z_0)}{n_k \rho(z_0)} + O(\|z - z_0\|^2), \end{aligned}$$

where the inequality (25) follows from the inequality $\rho(z) \geq (1/n_k)\ |\tau_k(z)|$ and the equation $\rho(z_0) = (1/n_k)\ |\tau_k(z_0)|$, and the equality (26) follows from the formula

$$\frac{1}{(a+b)} - \frac{1}{2b} = \frac{b-a}{2b(a+b)}$$

and the fact that $[|\tau_k(z_0)| - |\tau_k(z)|] = O(\|z - z_0\|)$. Therefore,

$$\begin{aligned}\rho(z) - \rho(z_0) \quad &- \quad \langle \frac{\lambda_k}{n_k \rho(z_0)} [\text{tr } \overline{B}_{k1}, \cdots, \ \text{tr } \overline{B}_{ks}], z - z_0 \rangle \\ &= \quad \rho(z) - \rho(z_0) - \frac{1}{n_k \rho(z_0)} \text{Re } (\overline{\lambda}_k (\nabla \tau_k(z_0)^T (z - z_0))) \\ &\geq \quad O(\|z - z_0\|^2)\end{aligned}$$

□

Although the results stated in Theorems 5 and 6 are significant, they do not clarify the status of the remaining vectors in the sets Φ, Ξ and Ψ. We have yet to establish that these vectors are proximal normal subgradients. Indeed, we are uncertain whether or not this is the case. The difficulty stems from the fact that those vectors involving the matrix N_k arise from the bifurcation of the eigenvalues into Puiseux-Newton series of the form (3) where the power p satisfies $(1/n) \leq p < 1$. Because of this, the individual behavior of these eigenvalues is not fully captured by the trace of their associated λ_k-group since these traces are analytic at z_0. However, it was the properties of these traces, especially their analyticity, that allowed us to establish the results in Theorems 5 and 6. Consequently, a different approach may be necessary in order to establish whether or not those vectors in the sets Φ, Ξ and Ψ involving the matrix N_k are proximal normal subgradients.

References

[1] V. I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys 26:2(1971), 29-43.

[2] H. Baumgärtel, *Analytic Perturbation Theory for Matrices and Operators*, Birkhäuser Verlag, Boston (1985).

[3] J. V. Burke and M. L. Overton, *On the subdifferentiability of functions of a matrix spectrum I: Mathematical foundations*, this volume.

[4] J. V. Burke and M. L. Overton, *Differential properties of eigenvalues*, in preparation.

[5] T. Kato, *Perturbation Theory for Linear Operators*, Grundlehren der mathematischen Wissenschaften 132, Springer-Verlag, New York (1980).

[6] B.S. Mordukhovich, *Approximation Methods in Optimization and Control Problems*, Moscow, Nauka, Main Physical and Mathematical Editions (1988) (In Russian).

[7] M. L. Overton and R. S. Womersley, *On minimizing the spectral radius of a nonsymmetric matrix function: Optimality conditions and duality theory*, SIAM J. Matrix Anal. Appl., 9(1988), 473-498.

[8] R.T. Rockafellar, *Extensions of Subgradient Calculus with Applications to Optimization*, Nonlinear Analysis, Theory, Methods and Applications, vol. 9, no. 7, 1985, pp.665–698.