



DIFFERENTIAL PROPERTIES OF THE SPECTRAL ABSCISSA AND THE SPECTRAL RADIUS FOR ANALYTIC MATRIX-VALUED MAPPINGS

JAMES V. BURKE†‡ and MICHAEL L. OVERTON§||

† Department of Mathematics, University of Washington, Seattle, WA 98195, U.S.A.; and
§ Computer Science Department, Courant Institute of Mathematical Sciences, New York University, New York, U.S.A.

(Received 28 August 1992; received in revised form 21 May 1993; received for publication 4 August 1993)

Key words and phrases: Matrix-valued mappings, differential properties.

1. INTRODUCTION

IN THIS study we consider the directional differentiability of two related functions of the spectrum of an analytic matrix function. Specifically, given an analytic (holomorphic) matrix-valued mapping $A(z)$ from \mathbb{C}^p to $\mathbb{C}^{n \times n}$, we are interested in the directional differentiability of the functions

$$\alpha(z) = \max\{\operatorname{Re} \lambda : \lambda \in \Sigma(z)\}, \tag{1}$$

and

$$\rho(z) = \max\{|\lambda| : \lambda \in \Sigma(z)\}, \tag{2}$$

where $\Sigma(z)$ is the spectrum of $A(z)$, i.e.

$$\Sigma(z) = \{\lambda : P(\lambda, z) = 0\}, \tag{3}$$

with

$$P(\lambda, z) = \det(\lambda I - A(z)), \tag{4}$$

the characteristic polynomial for $A(z)$. The elements of $\Sigma(z)$ are called eigenvalues and the functions α and ρ are called the spectral abscissa and spectral radius maps, respectively, for the analytic matrix function $A(z)$. These functions are of fundamental importance in many applications. Perhaps the most important of these is to the notion of stability for discrete and continuous dynamical systems.

Consider the differential equation

$$\dot{x} = A(z)x \tag{5}$$

for a fixed choice of the parameter $z \in \mathbb{C}^p$. Over 100 years ago, Lyapunov [1] showed that the solutions of this system attain a steady state (the trajectories stabilize) if and only if $\alpha(z) < 0$. As a consequence of this result the matrix $A(z)$ is said to be stable for a given value of z if $\alpha(z) < 0$. Although stable matrices have been intensely studied for the past two centuries, many fundamental properties have yet to be completely understood and so they remain a focal point of modern research [2, 3].

‡ The work of the author was supported in part by National Science Foundation Grant DMS-9102059.

|| The work of this author was supported in part by National Science Foundation Grant CCR-9101640.

The function ρ plays a similar role for the discrete system

$$x_{k+1} = A(z)x_k, \quad k = 0, 1, \dots \tag{6}$$

In this context, $A(z)$ is said to be stable if $\rho(z) < 1$. If $A(z)$ possesses this type of stability, then it is straightforward to show that the sequence $\{x_k\}$ converges [4].

A central theme in the study of stable matrices is the identification and characterization of those values of the *design* or *control* parameter z that yield stable solutions to the systems (5) and (6) [2, 3, 5]. The values of the parameter z that yield stability constitute the *stability region* for A . It is the desire to locate such values for the parameter z that provides the underlying motivation for this paper.

Observe that one can attempt to locate points in the stability region for A by minimizing α (respectively, ρ). For this purpose one is naturally led to an investigation of the variational properties of the mappings α and ρ subject to perturbations in z . The published literature on this subject contains very little information about these properties when the eigenvalues achieving the maximum value in the definition of either α or ρ have multiplicity greater than one. When the eigenvalues are *simple*, i.e. they have multiplicity one, then they are differentiable [6, theorem II.5.13a] and the variational properties of α and ρ are easily derived from the theory of max functions [7, 8]. If the eigenvalues achieving the max are *semisimple*, i.e. the corresponding part of the Jordan form of $A(z)$ is diagonal (equivalently, their associated factors in the polynomial of minimum degree annihilating $A(z)$ are linear), then the variational structure can be derived from the perturbation theory of linear operators [6, 8]. Beyond these two cases it appears that no further variational information is available. However, in practice, consideration of the general case of multiple eigenvalues is of the greatest importance, since the minimization of either α or ρ tends to cause the coalescence of eigenvalues. In the study of multiple eigenvalues, the most important case is not the semisimple case, but rather the nonderogatory case, i.e. the case where each eigenvalue corresponds to a single Jordan block in the Jordan form of $A(z)$ (equivalently, its multiplicity in both the characteristic and minimum polynomial for $A(z)$ is the same). In the absence of any assumption of special properties on the matrix such as symmetry, the nonderogatory case is generic [9].

In this study we consider the general case in which the eigenvalues may have any Jordan structure. In the case where one or more eigenvalues is defective, i.e. not semisimple, the functions α and ρ are, generally, non-Lipschitzian. Thus, nonstandard techniques are required in the analysis of their variational properties. In order to give some insight into the problem, consider the following two examples.

Example 1. Let

$$A(z) = \begin{bmatrix} 0 & 0 & z_1 + z_2^2 \\ 0 & 0 & 0 \\ z_1 + z_2^2 & 0 & 0 \end{bmatrix}$$

where for convenience we restrict z to real values. This matrix has a semisimple triple eigenvalue $\lambda = 0$ at $z = 0$. The eigenvalues of $A(z)$ are 0 and $\pm(z_1 + z_2^2)$, so

$$\alpha(z) = |z_1 + z_2^2|$$

(see Fig. 1).

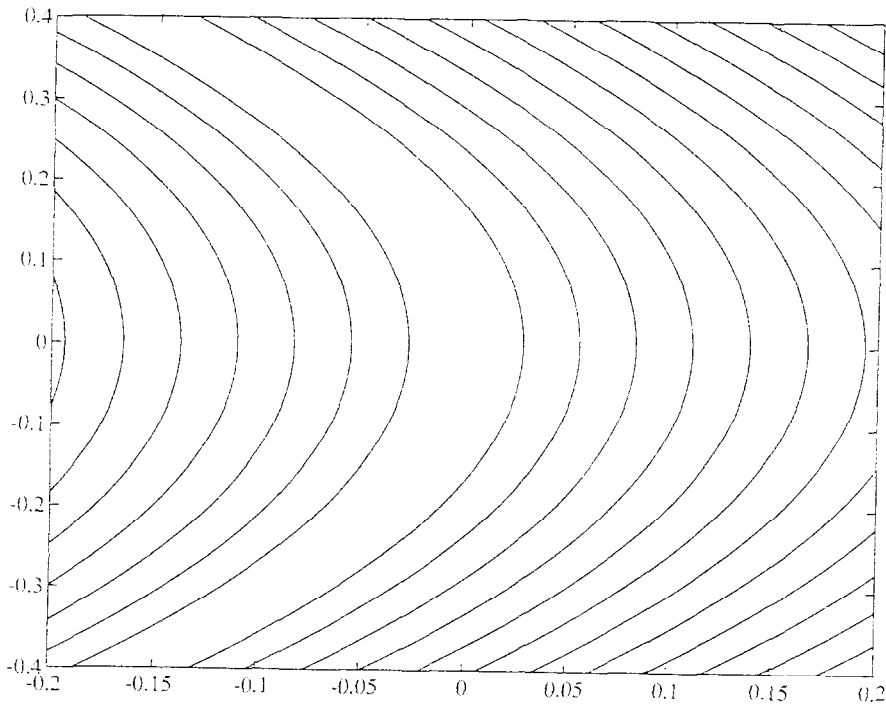


Fig. 1. Approximate contour lines for $\alpha(z)$ in example 1.

Example 2. Let

$$A(z) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ z_1 + z_2^2 & 0 & 0 \end{bmatrix}$$

where z is real. This matrix has a nonderogatory triple eigenvalue $\lambda = 0$ at $z = 0$. The eigenvalues of $A(z)$ are the three cube roots of $z_1 + z_2^2$, so

$$\alpha(z) = \kappa |z_1 + z_2^2|^{1/3}$$

where $\kappa = 1$ if $z_1 + z_2^2 \geq 0$, $\kappa = \frac{1}{2}$ otherwise (see Fig. 2).

Let us evaluate the ordinary directional derivative of α , defined by

$$\alpha'(z^0; d) = \lim_{\varepsilon \downarrow 0} \frac{\alpha(z^0 + \varepsilon d) - \alpha(z^0)}{\varepsilon} \tag{7}$$

for these two examples. We find that for example 1,

$$\alpha'(0; d) = \lim_{\varepsilon \downarrow 0} \frac{|\varepsilon d_1 + \varepsilon^2 d_2^2|}{\varepsilon} = |d_1|,$$

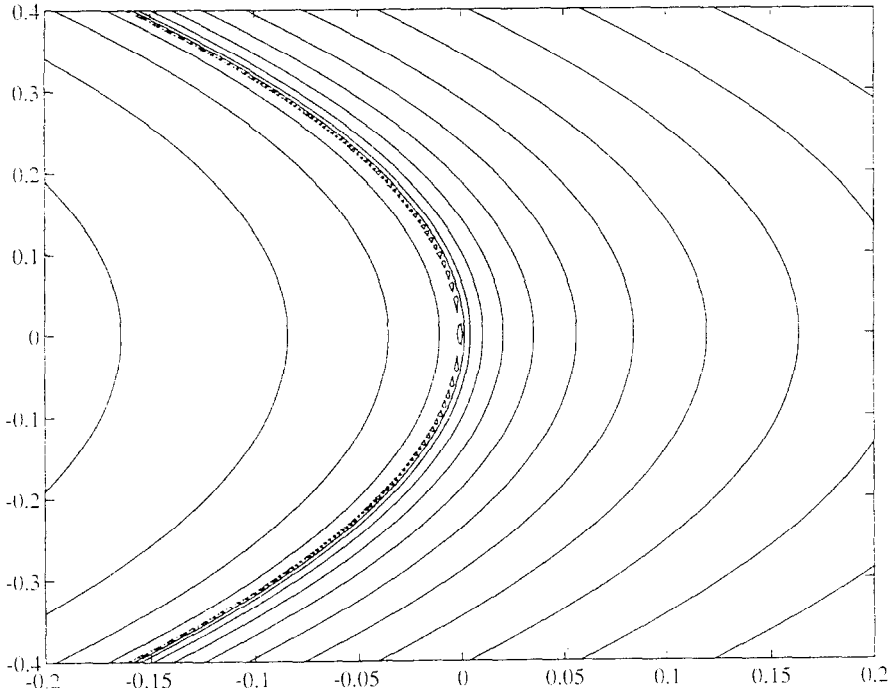


Fig. 2. Approximate contour lines for $\alpha(z)$ in example 2.

while for example 2,

$$\alpha'(0; d) = \lim_{\varepsilon \downarrow 0} \frac{|\varepsilon d_1 + \varepsilon^2 d_2^2|^{1/3}}{\varepsilon} = \infty$$

for every nonzero choice of d . These examples are typical: the ordinary directional derivative is always finite if the relevant eigenvalues achieving the spectral abscissa are semisimple, and almost always infinite otherwise. However, notice that in example 2, if $d = [0 \ 1]^T$, there does exist a *curve tangent to d* along which the difference quotient for the spectral abscissa is finite, in fact zero, namely $\gamma(\varepsilon) = [-\varepsilon^2 \ \varepsilon]^T$. We are, therefore, led to consider a directional derivative which depends on properties of α along curves tangent to the given direction.

Let $\gamma: \mathbb{C} \rightarrow \mathbb{C}^v$ be any analytic curve in \mathbb{C}^v and consider the eigenvalues of $\hat{A}(\zeta) = A(\gamma(\zeta))$ near $\zeta = 0$. It is well known that these eigenvalues, being the roots of the characteristic polynomial, are given by the Puiseux–Newton series, that is, series in fractional powers of ζ with the smallest such power being greater than or equal to $1/n$. The variational results that we obtain are a consequence of certain properties of these series. As we shall see, the derivation of these properties depends on a classical computational scheme due to Newton [10] known as the Puiseux–Newton diagram or the Newton Polygon [11].

This approach leads us to introduce a notion of directional differentiability that depends on analyticity in the following way: for $w: \mathbb{C}^v \rightarrow \mathbf{R}$, define $w^h(z; \cdot): \mathbb{C}^v \rightarrow \mathbf{R} \cup \{\pm\infty\}$ by

$$w^h(z; d) = \inf_{\gamma \in \Gamma(z, d)} \liminf_{\varepsilon \downarrow 0} \frac{w(\gamma(\varepsilon)) - w(z)}{\varepsilon}, \tag{8}$$

where

$$\Gamma(z, d) = \{\gamma: \mathbb{C} \rightarrow \mathbb{C}^r \mid \gamma \text{ is analytic, } \gamma(0) = z, \text{ and } \gamma'(0) = d\}. \quad (9)$$

The superscript h in (8) is used to emphasize that $w^h(z; d)$ depends only on the holomorphic curves in \mathbb{C}^r that pass through z . This notion of directional differentiability shares many of the properties of other such notions that have recently been developed in the literature on nonsmooth analysis [12–16]. Indeed, the directional derivative (8) is closely related to the lower Dini directional derivative and, consequently, to the notion of lower semigradients or contingent derivatives [13]. Understanding the relationship between (8) and these other types of directional differentiability is important for the development of a calculus. Based on the results of this article, partial results in this direction are presented in [17]. These results concern the relationship to the proximal subdifferential introduced by Clarke in [18]. However, a great deal more work needs to be done. We defer further discussion of these issues to future work. At present, we concentrate on the evaluation of (8) for the functions α and ρ . It is clear that for example 1,

$$\alpha^h(0; d) = \alpha'(0; d) = |d_1|$$

while for example 2,

$$\alpha^h(0; d) = 0$$

for d equal to a multiple of $[0 \ 1]^T$ and ∞ otherwise.

The paper is organized as follows. We begin in Section 2 by deriving a lower bound on $\alpha^h(z; d)$ in terms of the *coefficients of the characteristic polynomial*; this lower bound is achieved if a nondegeneracy condition is assumed. Then, in Section 3, we use the results of Section 2 to obtain a characterization of $\alpha^h(z; d)$ in terms of the matrix $A(z)$ itself, again establishing a general lower bound. The nondegeneracy condition required to establish that the lower bound is sharp imposes an important constraint on the eigenvalues of the matrix $A(z)$: specifically, the condition can hold only if the eigenvalues achieving the maximum value in (1) are nonderogatory. As already noted, this is the generic case. The opposite extreme, namely the least generic case, occurs when all the eigenvalues achieving the maximum value are semisimple. In this case, however, one can precisely evaluate the ordinary directional derivative. This is done in [8]. Thus, the cases for which precise results are unknown are those where at least one of the eigenvalues achieving the maximum value is both derogatory and defective. Nonetheless, the lower bounds that we establish do provide a great deal of information about when these directional derivatives attain the value $+\infty$.

Finally, in Section 4 we extend the results of Sections 2 and 3 to the case of the spectral radius function, obtaining a precise value for $\rho^h(z; d)$ in both the nonderogatory and semisimple cases. The spectral radius case is somewhat more complicated than the spectral abscissa case because of the singularity at the origin, but the results have the same essential character.

In the discussion that follows certain statements are sensitive to the domain of the variable being discussed. Thus, in order to avoid confusion, we obey the following convention concerning the labeling of the variables ζ , ε , and z : ζ will always represent a complex scalar, ε a real scalar, and z a vector in \mathbb{C}^r .

2. VARIATIONAL PROPERTIES OF THE SPECTRAL ABCISSA IN TERMS OF COEFFICIENTS OF THE CHARACTERISTIC POLYNOMIAL

Let $\mathfrak{C}_G[\lambda]$ be the set of monic polynomials in λ whose coefficients are analytic mappings from $G \subset \mathbb{C}^r$ to \mathbb{C} . If $G = \mathbb{C}^r$, we simply write $\mathfrak{C}[\lambda]$. Let P be the polynomial defined in (4). Clearly, P is contained in $\mathfrak{C}[\lambda]$. The results of this section only depend on this last fact. Thus, these results apply to any polynomial in $\mathfrak{C}[\lambda]$. From [19, pp. 376–381] there exists a neighborhood G of z^0 in \mathbb{C}^r on which P has the unique representation

$$P = \prod_{\lambda_k \in \Sigma(z^0)} \mu_k, \tag{10}$$

where, for each $\lambda_k \in \Sigma(z^0)$, the polynomial $\mu_k \in \mathfrak{C}_G[\lambda]$ takes the form

$$\mu_k(\lambda, z) = (\lambda - \lambda_k)^{t_k} + c_{k1}(z)(\lambda - \lambda_k)^{t_k-1} + \dots + c_{kt_k}(z), \tag{11}$$

where

$$c_{kj}(z^0) = 0 \quad \text{for } j = 1, \dots, t_k$$

and t_k is the multiplicity of the root λ_k . We begin with the case in which P has only a single root λ_0 , with multiplicity t_0 , i.e.

$$P(\lambda, z) = (\lambda - \lambda_0)^{t_0} + c_1(z)(\lambda - \lambda_0)^{t_0-1} + \dots + c_{t_0}(z), \tag{12}$$

where

$$c_j(z^0) = 0 \quad \text{for } j = 1, \dots, t_0.$$

We will return to the general case at the end of this section.

Our first objective is to understand the behavior of α along analytic curves in \mathbb{C}^r passing through z^0 . Let $\gamma: \mathbb{C} \rightarrow \mathbb{C}^r$ be an analytic curve satisfying $\gamma(0) = z^0$. Compose each c_j with γ to obtain

$$P(\lambda, \gamma(\zeta)) = (\lambda - \lambda_0)^{t_0} + \beta_1(\zeta)(\lambda - \lambda_0)^{t_0-1} + \dots + \beta_{t_0}(\zeta) = 0, \tag{13}$$

a polynomial equation in λ with analytic coefficients $\beta_j(\zeta) = c_j(\gamma(\zeta))$, satisfying

$$\beta_j(0) = 0, \quad j = 1, \dots, t_0.$$

We may write

$$\beta_j(\zeta) = \beta_j^{(1)}\zeta + \beta_j^{(2)}\zeta^2 + \dots$$

where, for example,

$$\beta_j^{(1)} = c_j'(z^0)\gamma'(0). \tag{14}$$

In the discussion which follows we restrict ζ to a nontrivial real interval $[0, \varepsilon_0]$ and write ε in place of ζ to emphasize this restriction.

As has already been noted, it is well known (e.g. [6, 19]) that the roots of (13) are described by series in fractional powers of ε . These series are commonly called Puiseux–Newton series, since it was Puiseux [20] who established their convergence; however, they were derived formally by Newton two centuries earlier (see also [11, Chapter 1, Section 2; 10, p. 88] for examples and applications). We obtain the results we need by making use of a diagram devised by Newton for the purpose of calculating coefficients of Puiseux–Newton series.

Let $\hat{\beta}_j = \beta_j^{(l_j)}$ be the first nonzero value in the sequence $\{\beta_j^{(1)}, \beta_j^{(2)}, \dots\}$. By definition, $l_j \geq 1$, $j = 1, \dots, t_0$. If $\beta_j(\varepsilon)$ is identically zero, take $l_j = \infty$; also, since the coefficient of $(\lambda - \lambda_0)^{t_0}$ in

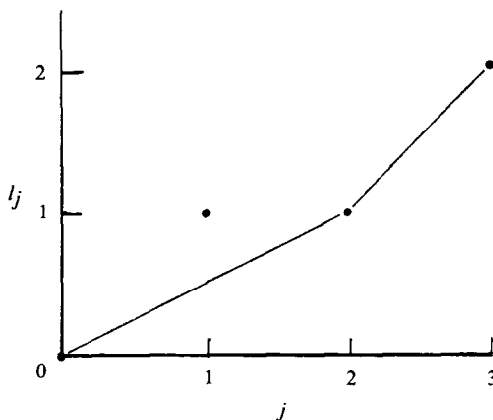


Fig. 3.

$P(\lambda, \varepsilon)$ is one, take $l_0 = 0, \hat{\beta}_0 = 1$. Now plot the values l_j vs j , and consider the lower boundary of the convex hull of the points plotted. Let s_j be the slope of the line on $[j, j + 1]$ forming part of this boundary, $j = 0, \dots, t_0 - 1$. Clearly $1/t_0 \leq s_0 \leq s_1 \leq \dots \leq s_{t_0-1}$. Figure 3 shows the diagram for the following example (taken from [19])

$$t_0 = 3; \quad \lambda_0 = 0; \quad \beta_1(\varepsilon) = \varepsilon; \quad \beta_2(\varepsilon) = -\varepsilon - \varepsilon^2; \quad \beta_3(\varepsilon) = \varepsilon^2 + 2\varepsilon^3.$$

We have $l_0 = 0, l_1 = 1, l_2 = 1, l_3 = 2$, and so $s_0 = s_1 = \frac{1}{2}, s_2 = 1$.

Now consider the following ‘‘Ansatz’’ argument. Suppose a root of (13) is to have the form

$$\lambda(\varepsilon) - \lambda_0 = a\varepsilon^p + \dots \tag{15}$$

where a is nonzero and p is the smallest power of ε in the expansion for this root. Substituting (15) into (13), we need

$$(a^{t_0}\varepsilon^{t_0p} + \dots) + (\hat{\beta}_1\varepsilon^{l_1} + \dots)(a^{t_0-1}\varepsilon^{(t_0-1)p} + \dots) + \dots + (\hat{\beta}_{t_0-1}\varepsilon^{l_{t_0-1}} + \dots)(a\varepsilon^p + \dots) + (\hat{\beta}_{t_0}\varepsilon^{l_{t_0}} + \dots) = 0.$$

The terms involving the smallest powers of ε are among the terms

$$a^{t_0}\varepsilon^{t_0p}, \hat{\beta}_1 a^{t_0-1}\varepsilon^{l_1+(t_0-1)p}, \dots, \hat{\beta}_{t_0-1} a\varepsilon^{l_{t_0-1}+p}, \hat{\beta}_{t_0}\varepsilon^{l_{t_0}}. \tag{16}$$

For cancellation to take place, at least two terms with the same smallest power of ε must appear. Equivalently, p must equal one or more of the slopes s_0, \dots, s_{t_0-1} defined by the Puiseux–Newton diagram. The following discussion will apply to a particular choice of such p . Define b and g by $p = s_b = \dots = s_{b+g-1}$, so that the line in the diagram with slope p passes from the point (b, l_b) to the point $(b + g, l_{b+g})$. Cancellation of the coefficients of the terms with the smallest powers of ε in (16) requires a to be the root of a polynomial equation with degree g , with leading term $\hat{\beta}_b a^g$ and constant term $\hat{\beta}_{b+g}$, and with an additional intermediate nonzero term for each point (j, l_j) lying on the line in the diagram with slope p , where $b < j < b + g$. Now let $p = q/f$, where q, f are relatively prime integers. By definition, p is an integral multiple of $1/g$, so g is an integral multiple of f , say $g = mf$. It is then clear from the diagram that of the $g - 1$ abscissa values j between b and $b + g$, only every f th value is a candidate for

the intersection of the line with a point with integer coordinates. Consequently the polynomial of degree g in a reduces to a polynomial of degree m in a^f , which we may denote by $Q(r)$. The conclusion is that the given value of p is associated with g roots with an expansion of the form (15), with a taking the values

$$r_h^{1/f} \omega^j, h = 1, \dots, m, \quad j = 1, \dots, f \tag{17}$$

where the r_h are the m roots of $Q(r) = 0$, $r_h^{1/f}$ is the principal f th root of r_h and ω is the principal f th root of unity.

Completing the example given above, we see that the two values for p are $s_0 = s_1 = \frac{1}{2}$ and $s_2 = 1$. In the case $p = \frac{1}{2}$ we have $b = 0, g = 2, f = 2, m = 1$, with $Q(r) = r - 1$, so the possible values for a are ± 1 , giving the Puiseux–Newton series

$$\lambda(\varepsilon) - \lambda_0 = \pm \varepsilon^{1/2} + \dots$$

In the case $p = 1$ we have $b = 2, g = 1, f = 1, m = 1$, with $Q(r) = r - 1$, so the only possible value for a is 1, giving the Puiseux–Newton series

$$\lambda(\varepsilon) - \lambda_0 = \varepsilon + \dots$$

The subsequent terms in the series may also be calculated by repeating the process.

In the case where the polynomial P in (10) arises as the characteristic polynomial of an analytic matrix-valued mapping, the Puiseux–Newton series that can occur have been completely characterized in the work of Lidskii [21] (see also [19, Section 7.4]), and Langer and Najman [22]. Lidskii employs a sophisticated approach predicated on his extension of Kato’s reduction technique [6]. Langer and Najman take a completely different tack. Their approach is based on the local Smith form (see [23, p. 331]). However, for the application at hand, such precision is not required. We provide an alternative development which is both elementary and self-contained. The key is provided by the next result. It is a generalization of [24, theorem 1]; see also [25]. The lemma gives information about the coefficients of (13) when it is assumed that roots of (13) lie within $O(\varepsilon)$ of the half plane $\text{Re } \bar{y}_0(\zeta - \lambda_0) \leq 0$.

LEMMA 1. Consider the polynomial equation (13), with roots given by one or more Puiseux–Newton series of the form (15). Let $y_0 \in \mathbb{C}$ and suppose that there exists $\varepsilon_0 > 0$ such that all the roots $\lambda(\varepsilon)$ of (13) satisfy

$$\text{Re } \bar{y}_0(\lambda(\varepsilon) - \lambda_0) \leq \delta\varepsilon + o(\varepsilon). \tag{18}$$

Then

$$\text{Re } \bar{y}_0 \beta_1^{(1)} \geq -t_0 \delta, \tag{19}$$

$$\text{Re } \bar{y}_0^2 \beta_2^{(1)} \geq 0, \quad \text{Im } \bar{y}_0^2 \beta_2^{(1)} = 0, \tag{20}$$

$$\beta_j^{(1)} = 0, \quad j = 3, \dots, t_0. \tag{21}$$

Here (20) is understood to be vacuous if $t_0 = 1$.

Proof. The coefficient $\beta_1(\varepsilon)$ is the sum of the differences $\lambda_0 - \lambda(\varepsilon)$ over the roots $\lambda(\varepsilon)$ of (13); thus, (19) follows from (18), letting $\varepsilon \rightarrow 0$. The other results follow from the Puiseux–Newton diagram as follows. Consider the Puiseux–Newton series corresponding to $p = s_0$, the smallest

possible value. In order for (18) to hold either:

- (i) $p \geq 1$ (e.g. if $f = 1$); or
- (ii) $p = \frac{1}{2}$, $f = 2$, and $\text{Re } \bar{y}_0 r_h^{1/2} = 0$ for $h = 1, \dots, m$, where the r_h are the m roots of $Q(r)$, with $Q(r)$ taking the form

$$Q(r) = r^m + \beta_2^{(1)} r^{m-1} + \dots + \beta_{2m}^{(m)}. \tag{22}$$

No other cases having $p < 1$ are possible due to the splitting of the roots as described in (17). In both cases (i) and (ii), $p > \frac{1}{2}$, so $\beta_j^{(1)} = 0$ for $j = 3, \dots, t_0$ from the Puiseux–Newton diagram. In the case $p \geq 1$, we also have $\beta_2^{(1)} = 0$. In the case $f = 2$, observe that the condition $\text{Re } \bar{y}_0 r_h^{1/2} = 0$ is equivalent to the two conditions $\text{Re } \bar{y}_0^2 r_h \leq 0$ and $\text{Im } \bar{y}_0^2 r_h = 0$. Now, since $-\beta_2^{(1)}$ is the sum of the roots of $Q(r)$, (20) follows. ■

We now apply this result to the evaluation of $\alpha^h(z^0; d)$, defined by (1), (3) and (8), where P is assumed to have the form (12). First observe that we can replace the limit infimum in (8) by limit since the perturbed roots of the polynomial (13) are given by Puiseux–Newton series of the form (15) for some nonnegative rational number p . Therefore, this limit always exists and can only take the value $+\infty$ if it is not finite. Consequently, $\alpha^h(z^0; \cdot): \mathbb{C}^\nu \mapsto \mathbf{R} \cup \{+\infty\}$ and is given by

$$\alpha^h(z; d) = \inf_{\gamma \in \Gamma(z, d)} \lim_{\varepsilon \downarrow 0} \frac{\alpha(\gamma(\varepsilon)) - \alpha(z)}{\varepsilon}. \tag{23}$$

We now state the main result of this section.

THEOREM 2. Define α by (1) and (3) where P has the form (12) near $z^0 \in \mathbb{C}^\nu$ and choose $d \in \mathbb{C}^\nu$. If any one of the conditions

$$\text{Re } c'_2(z^0)d \geq 0, \quad \text{Im } c'_2(z^0)d = 0, \tag{24}$$

$$c'_j(z^0)d = 0, \quad j = 3, \dots, t_0 \tag{25}$$

is violated, then

$$\alpha^h(z^0; d) = +\infty;$$

otherwise

$$\alpha^h(z^0; d) \geq -\frac{1}{t_0} \text{Re } c'_1(z^0)d. \tag{26}$$

Moreover, if the rank of $c'(z^0)$ is t_0 , where $c: \mathbb{C}^\nu \mapsto \mathbb{C}^{t_0}$ is given by

$$c(z) = \begin{bmatrix} c_1(z) \\ \vdots \\ c_{t_0}(z) \end{bmatrix},$$

then equality holds in (26) whenever (24) and (25) are satisfied.

Proof. Suppose $+\infty > \delta > \alpha^h(z^0; d)$. Then there is a $\gamma \in \Gamma(z^0, d)$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{\alpha(\gamma(\varepsilon)) - \alpha(z^0)}{\varepsilon} < \delta, \tag{27}$$

or equivalently,

$$\alpha(\gamma(\varepsilon)) - \alpha(z^0) < \delta\varepsilon \quad \text{for } \varepsilon \in [0, \varepsilon_0],$$

for some $\varepsilon_0 > 0$. Let $\beta_j^{(1)} = c'_j(z^0)d$ for $j = 1, 2, \dots, t_0$ as in (14). By invoking lemma 1 with $y_0 = 1$, we see that (24) and (25) must be satisfied. Thus, if any one of these conditions is violated, we must have $\alpha^h(z^0; d) = +\infty$. By letting $\delta \downarrow \alpha^h(z^0; d)$, we also obtain from lemma 1 the inequality

$$\alpha^h(z^0; d) \geq -\frac{1}{t_0} \operatorname{Re} \beta_1^{(1)}.$$

Let us now suppose that the rank of $c'(z^0)$ is t_0 . We need to establish equality in (26). Clearly, we need only consider the case in which (24) and (25) hold. In this case, equality follows if we can exhibit a curve $\gamma \in \Gamma(z^0, d)$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{\alpha(\gamma(\varepsilon)) - \alpha(z^0)}{\varepsilon} = -\frac{1}{t_0} \operatorname{Re} \beta_1^{(1)}. \tag{28}$$

Consider the coefficients of the powers of $(\lambda - \lambda_0)$ in the polynomial

$$\begin{aligned} & \left((\lambda - \lambda_0) + \frac{\beta_1^{(1)}}{t_0} \zeta \right)^{t_0-2} \left(\lambda - \lambda_0 + i(\beta_2^{(1)}\zeta)^{1/2} + \frac{\beta_1^{(1)}}{t_0} \zeta \right) \left(\lambda - \lambda_0 - i(\beta_2^{(1)}\zeta)^{1/2} + \frac{\beta_1^{(1)}}{t_0} \zeta \right) \\ &= (\lambda - \lambda_0)^{t_0} + \beta_1^{(1)}\zeta(\lambda - \lambda_0)^{t_0-1} + (\beta_2^{(1)}\zeta + O(\zeta^2))(\lambda - \lambda_0)^{t_0-2} + \dots, \\ &\equiv (\lambda - \lambda_0)^{t_0} + v_1(\zeta)(\lambda - \lambda_0)^{t_0-1} + v_2(\zeta)(\lambda - \lambda_0)^{t_0-2} + \dots, \end{aligned} \tag{29}$$

where $i = \sqrt{-1}$. Note that these coefficients satisfy (20) and (21) with $y_0 = 1$. Also note that if γ can be chosen from (9) so that (13) has these coefficients, then (28) is satisfied for this curve and the proof is complete. We now show that this can indeed be done.

Define $F: \mathbb{C}^{v+1} \rightarrow \mathbb{C}^{t_0}$ by

$$F(z, \zeta) = c(z) - v(\zeta),$$

where $v: \mathbb{C} \rightarrow \mathbb{C}^{t_0}$ is the curve whose component functions are the coefficients of the powers of $(\lambda - \lambda_0)$ in the polynomial (29)

$$v(\zeta) = [v_1(\zeta), v_2(\zeta), \dots, v_t(\zeta)]^T = \zeta c'(z^0)\gamma'(0) + O(\zeta^2), \tag{30}$$

where the second equality follows from the definition of v . Let $I \subset \{1, \dots, v\}$ be such that the matrix $c'_I(z^0)$ is nonsingular and set $J = \{1, \dots, v\} \setminus I$. By the implicit function theorem [26], there is a neighborhood $\bar{G} \subset \mathbb{C}^{v-t_0+1}$ of $((z^0)_J, 0)$ and an analytic mapping $\hat{\gamma}: \bar{G} \rightarrow \mathbb{C}^{t_0}$ such that

$$F(\hat{\gamma}(z_J, \zeta), z_J, \zeta) = 0,$$

for all $(z_J, \zeta) \in \bar{G}$ with

$$\hat{\gamma}((z^0)_J, 0) = (z^0)_I.$$

Furthermore,

$$\hat{\gamma}'((z^0)_J, 0) = -c'_I(z^0)^{-1}[c'_J(z^0), -v'(0)]. \tag{31}$$

Define $\gamma: \mathbb{C} \rightarrow \mathbb{C}^v$ by

$$\gamma_J(\zeta) = (z^0)_J + \zeta d_J, \tag{32}$$

and

$$\gamma_I(\zeta) = \hat{\gamma}(\gamma_J(\zeta), \zeta). \tag{33}$$

Then for all ζ sufficiently small $c(\gamma(\zeta)) = v(\zeta)$. Consequently, with this choice of γ , the polynomial (13) is precisely the polynomial (29). Moreover, from (30)–(33) we have

$$\gamma'_J(0) = d_J$$

and

$$\gamma'_I(0) = \hat{\gamma}'(\gamma_J(0), 0) \begin{bmatrix} d_J \\ 1 \end{bmatrix} = d_I,$$

so that $\gamma'(0) = d$. Thus, $\gamma \in \Gamma(z^0, d)$. This concludes the proof of the theorem. ■

Let us now return to the general case and recall the factorization (10). Our results in this case follow directly from the special case (12).

THEOREM 3. Let P have the general form (10) near $z^0 \in \mathbb{C}^{\nu}$, with α defined by (1) and (3) as before. Set

$$\mathcal{Q}_0(z^0) = \{\lambda_k \in \Sigma(z^0): \operatorname{Re} \lambda_k = \alpha(z^0)\}, \tag{34}$$

and choose $d \in \mathbb{C}^{\nu}$. If for some $\lambda_k \in \mathcal{Q}_0(z^0)$ any one of the conditions

$$\operatorname{Re} c'_{k2}(z^0)d \geq 0, \quad \operatorname{Im} c'_{k2}(z^0)d = 0, \tag{35}$$

$$c'_{kj}(z^0)d = 0, \quad j = 3, \dots, t_k, \tag{36}$$

is violated, then

$$\alpha^h(z^0; d) = +\infty; \tag{37}$$

otherwise,

$$\alpha^h(z^0; d) \geq \max \left\{ -\frac{\operatorname{Re} c'_{k1}(z^0)d}{t_k} : \lambda_k \in \mathcal{Q}_0(z^0) \right\}. \tag{38}$$

Moreover, if the vectors $\{c'_{kj}(z^0): \lambda_k \in \mathcal{Q}_1(z^0, d), j = 1, \dots, t_k\}$ are linearly independent, where $\mathcal{Q}_1(z^0, d)$ is the set

$$\left\{ \lambda_k \in \mathcal{Q}_0(z^0): \frac{\operatorname{Re} c'_{k1}(z^0)d}{t_k} = \min \left\{ \frac{\operatorname{Re} c'_{l1}(z^0)d}{t_l} : \lambda_l \in \mathcal{Q}_0(z^0) \right\} \right\}, \tag{39}$$

then equality holds in (38) when (35) and (36) hold.

Proof. The proof is almost identical to that of theorem 2. The primary difference is that now all of the roots in $\mathcal{Q}_0(z^0)$ contribute to the value of $\alpha^h(z^0; d)$. In order to see this observe that the inequality (27) implies that

$$\operatorname{Re}(\lambda(\gamma(\varepsilon)) - \lambda_k) < \delta\varepsilon \quad \text{for } \varepsilon \in [0, \varepsilon_0] \text{ and } \lambda_k \in \mathcal{Q}_0(z^0),$$

for some $\varepsilon_0 > 0$. Consequently, (38) again follows from lemma 1.

In order to establish equality in (38), it is again sufficient to exhibit the existence of a curve $\gamma \in \Gamma(z^0, d)$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{\alpha(\gamma(\varepsilon)) - \alpha(z)}{\varepsilon} = \max \left\{ -\frac{\operatorname{Re} c'_{k1}(z^0)d}{t_k} : \lambda_k \in \mathcal{G}_0(z^0) \right\}. \tag{40}$$

We generate such a curve precisely as in the proof of theorem 2 except now we must choose the curve from $\Gamma(z^0, d)$ so as to match the coefficients in (29) for each $\lambda_k \in \mathcal{G}_1(z^0, d)$. Just as before, it is the linear independence of the gradients $\{c'_{k1}(z^0): \lambda_k \in \mathcal{G}_1(z^0, d)\}$ which guarantees that this can be done via the implicit function theorem. Moreover, it is clear that we need only match the coefficients for $\lambda_k \in \mathcal{G}_1(z^0, d)$ since these are the dominant first order terms. ■

3. VARIATIONAL PROPERTIES OF THE SPECTRAL ABCISSA IN TERMS OF MATRIX ELEMENTS

Let $\mathcal{H}[\mathbb{C}^p, \mathbb{C}^{n \times n}]$ denote the set of mappings from \mathbb{C}^p to $\mathbb{C}^{n \times n}$ each of whose components is an analytic map from \mathbb{C}^p to \mathbb{C} . Let $A \in \mathcal{H}[\mathbb{C}^p, \mathbb{C}^{n \times n}]$. In this section we study the differential properties of the spectral abscissa α , defined by (1), (3), and (4). One can apply the results of the previous section to obtain differential information about α since P , defined in (4), is an element of $\mathcal{H}[\lambda]$. However, by itself, this result is not complete since it does not describe the relationship between A and the terms $c'_{kj}(z^0)d$ appearing in (38). In this section we describe this relationship, making use of results from [24] which in turn depend on work of Arnold [9]. Since our description depends on the Jordan decomposition of $A^{(0)} = A(z^0)$, we need to introduce the notation necessary for this discussion.

Suppose $A^{(0)}$ is a matrix with eigenvalues $\lambda_1, \dots, \lambda_\eta$, having multiplicities t_1, \dots, t_η , respectively. Let the Jordan form of $A^{(0)}$ be given by

$$A^{(0)} = SJS^{-1}$$

where

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_\eta \end{bmatrix},$$

$$J_k = \begin{bmatrix} J_{k1} & & \\ & \ddots & \\ & & J_{km_k} \end{bmatrix},$$

and the Jordan block

$$J_{kl} = \begin{bmatrix} \lambda_k & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot & 1 \\ & & & & & \lambda_k \end{bmatrix}$$

has dimension n_{kl} . We have

$$n_{k1} + \dots + n_{km_k} = t_k, \quad k = 1, \dots, \eta.$$

If $m_k = 1$, λ_k is said to be a nonderogatory eigenvalue, while if $m_k = t_k$, i.e. $n_{k1} = \dots = n_{km_k} = 1$, λ_k is said to be semisimple (nondefective).

Definition 4. Define the j th generalized trace of a square matrix A , denoted by

$$\text{tr}^{(j)}A,$$

as the sum of the elements on the diagonal of A which is $j - 1$ positions below the main diagonal. Thus, one obtains the ordinary trace in the case $j = 1$ and the bottom left element of the matrix in the case that j is the dimension of the matrix. If j exceeds the dimension of A , take $\text{tr}^{(j)}A = 0$.

THEOREM 5. Let $A \in \mathcal{JC}[\mathbb{C}^v, \mathbb{C}^{n \times n}]$ and choose $z^0 \in \mathbb{C}^v$. Suppose that $A^{(0)} = A(z^0)$ has Jordan form as described above. Define

$$A_q^{(1)} = \frac{\partial A}{\partial z_q}(z^0), \quad \text{for } q = 1, \dots, v.$$

For each $q = 1, \dots, v$ partition $S^{-1}A_q^{(1)}S$ conformally with the partition of J and denote its diagonal block corresponding to J_k by B_{qk} , $k = 1, \dots, \eta$, with each B_{qk} having diagonal blocks B_{qkl} corresponding to J_{kl} , $l = 1, \dots, m_k$. Then, for $k = 1, \dots, \eta$,

$$c'_{kj}(z^0) = \begin{bmatrix} -\sum_{l=1}^{m_k} \text{tr}^{(j)}B_{1kl} \\ \vdots \\ -\sum_{l=1}^{m_k} \text{tr}^{(j)}B_{vkl} \end{bmatrix}, \quad \text{for } j = 1, \dots, t_k, \tag{41}$$

where the functions c_{kj} are as given in (11) for the factorization (10) of $P(\lambda, z) = \det[\lambda I - A(z)]$ at $z = z^0$.

Proof. Let $d \in \mathbb{C}^v$ and $\gamma \in \Gamma(z^0, d)$ and define $\hat{A} = A \circ \gamma$. Given $M \in \mathbb{C}^{n \times n}$ denote by M_{kl} that block of M which conforms to the block J_{kl} of J . By [24, theorem 4] we have

$$\begin{aligned} -c'_{kj}(z^0)d &= \sum_{l=1}^{m_k} \text{tr}^{(j)}(S^{-1}\hat{A}'(0)S)_{kl} \\ &= \sum_{l=1}^{m_k} \text{tr}^{(j)}(S^{-1}A'(0)dS)_{kl} \\ &= \sum_{l=1}^{m_k} \text{tr}^{(j)}\left(S^{-1}\left(\sum_{q=1}^v d_q A_q^{(1)}\right)S\right)_{kl} \\ &= \sum_{q=1}^v d_q \sum_{l=1}^{m_k} \text{tr}^{(j)}B_{qkl}, \end{aligned}$$

for $j = 1, \dots, t_k$. Since this holds for all $d \in \mathbb{C}^v$, the result follows. ■

Remark. A generalization of this result, based on the *block diagonal form* of A , is given in [27, lemma 3.2]. The block diagonal form of a matrix is not unique and includes the Jordan form

as a special case [28, Section 7.1.3]. From the computational point of view, this approach has certain advantages since one need not compute the Jordan form in order to evaluate the derivatives c'_{kj} .

Thus, given the Jordan form (or, block diagonal form) of $A(z^0)$ together with $A'(z^0)$ it is possible to use theorems 3 and 5 to evaluate a lower bound for $\alpha^h(z^0; d)$. This result is formally stated in the following theorem.

THEOREM 6. Let the assumptions of theorem 5 hold, and define α by (1), (3), and (4). Choose $d \in \mathbb{C}^v$, and define $\mathcal{Q}_0(z^0)$ and $\mathcal{Q}_1(z^0, d)$ by (34) and (39), respectively. If for some $\lambda_k \in \mathcal{Q}_0(z^0)$ any one of the conditions

$$\begin{aligned} \operatorname{Re} \sum_{q=1}^v d_q \sum_{i=1}^{m_k} \operatorname{tr}^{(2)} B_{qki} &\leq 0, & \operatorname{Im} \sum_{q=1}^v d_q \sum_{l=1}^{m_k} \operatorname{tr}^{(2)} B_{qkl} &= 0, \\ \sum_{q=1}^v d_q \sum_{l=1}^{m_k} \operatorname{tr}^{(j)} B_{qkl} &= 0, & j &= 3, \dots, t_k, \end{aligned}$$

is violated, then

$$\alpha^h(z^0; d) = +\infty; \tag{42}$$

otherwise,

$$\alpha^h(z^0; d) \geq \max \left\{ \frac{\operatorname{Re} \sum_{q=1}^v d_q \sum_{l=1}^{m_k} \operatorname{tr}^{(1)} B_{qkl}}{t_k} \mid \lambda_k \in \mathcal{Q}_0(z^0) \right\}. \tag{43}$$

Moreover, if the vectors

$$\begin{bmatrix} \sum_{l=1}^{m_k} \operatorname{tr}^{(j)} B_{1kl} \\ \vdots \\ \sum_{l=1}^{m_k} \operatorname{tr}^{(j)} B_{vkl} \end{bmatrix}, \quad \text{for } \lambda_k \in \mathcal{Q}_1(z^0, d) \text{ and } j = 1, \dots, t_k, \tag{44}$$

are linearly independent, then equality holds in (43).

Proof. The result is an immediate consequence of theorems 3 and 5. Note that, in the case that the linear independence condition holds, the implicit function theorem used in the proof of theorem 3 provides the proof of existence of a curve $\gamma(\zeta)$ along which the difference quotient limit for the spectral abscissa of the matrix A achieves the given lower bound. Thus, an argument such as given in [24, theorem 6] is not needed. ■

Remark. Note that $\sum_{l=1}^{m_k} \operatorname{tr}^{(j)} B_{qkl} = \operatorname{tr}^{(j)} B_{qk}$ when $j = 1$, but not when $j > 1$. This observation allows one to simplify the formula on the right-hand side of (43). However, in order to keep the notation consistent throughout the statement of theorem 6, we have not included this simplification into its statement.

It will be helpful to consider a special case. Suppose that $\mathcal{Q}_0(z^0)$ consists of a single eigenvalue, say λ_1 , with multiplicity t_1 . If this eigenvalue is nonderogatory, it is associated

with only one Jordan block J_1 , with dimension t_1 . Let $C = \sum_{q=1}^{\nu} d_q B_{q1}$. The necessary conditions for $\alpha^h(z^0; d)$ to be finite then reduce to

$$\operatorname{Re} \operatorname{tr}^{(2)} C \leq 0, \quad \operatorname{Im} \operatorname{tr}^{(2)} C = 0, \quad \operatorname{tr}^{(j)} C = 0, \quad j = 3, \dots, t_1,$$

and its lower bound on the right-hand side of (43) reduces to

$$\frac{1}{t_1} \operatorname{Re} \operatorname{tr}^{(1)} C, \tag{45}$$

the average value of the real parts of the eigenvalues of C . In the specific case of example 2, Section 1, we obtain

$$C = \sum_{q=1}^2 d_q \frac{\partial A}{\partial z_q}(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ d_1 & 0 & 0 \end{bmatrix},$$

so the necessary condition for $\alpha^h(0; d)$ to be finite is $d_1 = 0$, and the lower bound is 0. On the other hand, if λ_1 is semisimple, it is associated with t_1 blocks J_{11}, \dots, J_{1t_1} , each of dimension one. In this case, the necessary conditions for α^h to be finite hold vacuously, and its lower bound on the right-hand side of (43) *again* reduces to (45). In the specific case of example 1, Section 1 we have

$$C = \sum_{q=1}^2 d_q \frac{\partial A}{\partial z_q}(0) = \begin{bmatrix} 0 & 0 & d_1 \\ 0 & 0 & 0 \\ d_1 & 0 & 0 \end{bmatrix}$$

so the lower bound on $\alpha^h(0; d)$ is 0.

It should be observed that if any eigenvalue $\lambda_k \in \mathcal{G}_1(z^0, d)$ is derogatory, then the vectors in (44) *cannot* be linearly independent. In order to see this, note that, for at least one j between 1 and t_k , j exceeds the dimension of all the blocks J_{kl} making up J_k , and, hence, the corresponding vector in (44) is zero. Thus, if $\mathcal{G}_1(z^0, d)$ contains a derogatory eigenvalue, then the sufficiency condition of theorem 3 is not satisfied. In this case, one should not expect to obtain equality in (43). Indeed, in the case where $\mathcal{G}_0(z^0)$ contains only semisimple eigenvalues, the resolvent theory for eigenvalue perturbations yields the following result.

THEOREM 7. If $\mathcal{G}_0(z^0)$ contains only semisimple eigenvalues, then for every $d \in \mathbb{C}^{\nu}$,

$$\alpha^h(z^0; d) = \alpha'(z^0; d) = \max_{\lambda_k \in \mathcal{G}_0(z^0)} \max_{1 \leq l \leq t_k} \operatorname{Re} \lambda'_{kl}, \tag{46}$$

where λ'_{kl} , $l = 1, \dots, t_k$ are the eigenvalues of $\sum_{q=1}^{\nu} d_q B_{qk}$ and $\alpha'(z^0; d)$ is the ordinary directional derivative defined in (7).

Proof. This is a consequence of [6, Section II.2.3, (2.40)], which shows that the eigenvalues of $A(\gamma(\zeta))$ corresponding to $\lambda_k \in \mathcal{G}_0(z^0)$ have the form

$$\lambda_k + \zeta \lambda'_{kl} + o(\zeta), \quad l = 1, \dots, t_k$$

regardless of which curve γ is chosen from $\Gamma(z^0, d)$. The matrix called P in [6] is the eigenprojection for λ_k and equals YZ , where Z consists of the columns of S corresponding to λ_k , which are right eigenvectors for λ_k , and Y consists of the corresponding rows of S^{-1} , which are left eigenvectors. For further discussion see [8, Section 3]. ■

Suppose $\mathcal{R}_0(z^0) = \{\lambda_1\}$. In the generic nonderogatory case, $\alpha^h(z^0; d)$ can be expected to achieve its lower bound, namely the *average* of the real parts of the eigenvalues of $C = \sum_{q=1}^v d_q B_{q1}$. On the other hand, theorem 7 shows that in the semisimple case, $\alpha^h(z^0; d)$ is the *maximum* of these quantities. Thus, considering example 1, Section 1 again we see that the lower bound on $\alpha^h(0; d)$ is 0, the average of the eigenvalues of C , but the actual value of this derivative is $|d_1|$, the maximum eigenvalue of C .

4. VARIATIONAL PROPERTIES OF THE SPECTRAL RADIUS

Let us now shift our attention to the study of the function ρ defined in (2). We study the differential properties of ρ in the same manner as we studied these properties for α . That is, we first consider the case when the spectrum near $z^0 \in \mathbb{C}^v$ is given as the set of roots of a polynomial of the form (12) and then extend this result to the general case. To this end, consider the directional derivative $\rho^h(z^0; d)$ defined by (8). As was the case for the directional derivative $\alpha^h(x; d)$, the limit infimum in the definition of $\rho^h(z^0; d)$ can be replaced by limit. This is justified in the same way as it was for $\alpha^h(z^0; d)$, that is, by considering the splitting behavior of the eigenvalues under perturbation. Thus, we may write

$$\rho^h(z; d) = \inf_{\gamma \in \Gamma(z^0, d)} \lim_{\varepsilon \downarrow 0} \frac{\rho(\gamma(\varepsilon)) - \rho(z)}{\varepsilon} \tag{47}$$

where $\Gamma(z^0, d)$ is defined in (9), and $\rho^h(z^0; \cdot): \mathbb{C}^v \mapsto \mathbf{R} \cup \{+\infty\}$. Continuing as in Section 2, we begin with the following key result for the case in which (12) holds.

THEOREM 8. Define ρ by (2) and (3) where P is given by (12) near $z^0 \in \mathbb{C}^v$ and choose $d \in \mathbb{C}^v$. We will consider two cases: $\rho(z^0) = 0$ and $\rho(z^0) \neq 0$.

(1) Assume that $\rho(z^0) = 0$ so that $\lambda_0 = 0$. In this case we have

$$\rho^h(z^0; d) \geq \begin{cases} \frac{1}{t_0} |c'_1(z^0)d|, & \text{if } c'_j(z^0)d = 0 \text{ for } j = 2, \dots, t_0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{48}$$

(2) Assume that $\rho(z^0) \neq 0$ and consider the conditions

$$\text{Re } \bar{\lambda}_0^2 c'_2(z^0)d \geq 0, \quad \text{Im } \bar{\lambda}_0^2 c'_2(z^0)d = 0, \tag{49}$$

$$c'_j(z^0)d = 0, \quad j = 3, \dots, t_0. \tag{50}$$

Then

$$\rho^h(z^0; d) \geq \begin{cases} \frac{1}{t_0 \rho(z^0)} [|c'_2(z^0)d| - \text{Re } \bar{\lambda}_0 c'_1(z^0)d], & \text{if (49) and (50) hold,} \\ +\infty, & \text{otherwise.} \end{cases} \tag{51}$$

Here it is understood that the function c_2 is identically zero if $t_0 = 1$. Moreover, if the rank of $c'(z^0)$ is t_0 , where $c: \mathbb{C}^r \rightarrow \mathbb{C}^{t_0}$ is given by

$$c(z) = \begin{bmatrix} c_1(z) \\ \vdots \\ c_t(z) \end{bmatrix},$$

then equality holds in (48) or (51) depending on whether $\rho(z^0) = 0$ or $\rho(z^0) \neq 0$ holds, respectively.

Proof. In this proof we will continue to use the notation of Section 2. Let $\gamma \in \Gamma(z^0, d)$, set $\beta_j^{(1)} = c_j'(z^0)\gamma'(0) = c_j'(z^0)d$ for $j = 1, 2, \dots, t_0$ as in (14), and let $\lambda(\varepsilon)$ be one of the roots (15) of (13). Since

$$|\lambda(\varepsilon)|^2 - |\lambda_0|^2 = 2 \operatorname{Re} \bar{\lambda}_0(\lambda(\varepsilon) - \lambda_0) + |\lambda(\varepsilon) - \lambda_0|^2, \tag{52}$$

a necessary condition for

$$\delta\varepsilon + o(\varepsilon) \geq |\lambda(\varepsilon)| - |\lambda_0|, \tag{53}$$

or equivalently,

$$2\delta|\lambda_0|\varepsilon + o(\varepsilon) \geq |\lambda(\varepsilon)|^2 - |\lambda_0|^2, \tag{54}$$

is that

$$\delta|\lambda_0|\varepsilon + o(\varepsilon) \geq \operatorname{Re} \bar{\lambda}_0(\lambda(\varepsilon) - \lambda_0). \tag{55}$$

It follows from lemma 1 that if either (20) or (21) with $y_0 = \lambda_0$, or equivalently, (49) or (50), do not hold, then inequality (54) cannot hold for any $\delta \in \mathbf{R}$. Since this is independent of $\gamma \in \Gamma(z^0, d)$, $\rho^h(z^0; d) = +\infty$ if any one of (49) or (50) are violated regardless of the value of $\rho(z^0)$.

Let us now suppose that (49) and (50) hold, i.e. (20) and (21) hold with $y_0 = \lambda_0$. Then for every $\delta > \rho^h(z^0; d)$ there is a $\gamma \in \Gamma(z^0, d)$ such that

$$\lim_{\varepsilon \downarrow 0} \frac{\rho(\gamma(\varepsilon)) - \rho(z^0)}{\varepsilon} < \delta, \tag{56}$$

or equivalently,

$$\rho(\gamma(\varepsilon)) - \rho(z^0) < \delta\varepsilon \quad \text{for } \varepsilon \in [0, \varepsilon_0],$$

for some $\varepsilon_0 > 0$. Therefore, inequalities (54) and (55) are satisfied.

As observed in Section 2, the roots of the equation (13) are necessarily Puiseux–Newton series of the form (15). Lemma 1 and inequality (55) imply that either:

- (i) the exponent p is greater than or equal to 1 corresponding to series of the form

$$\lambda(\varepsilon) = \lambda_0 + a\varepsilon + o(\varepsilon),$$

with $a \in \mathbb{C}$ possibly taking the value zero; or

- (ii) the exponent p equals $\frac{1}{2}$ corresponding to m pairs of roots of the form

$$\lambda(\varepsilon) = \lambda_0 \pm (r_h \varepsilon)^{1/2} + (\tilde{a}_\pm)\varepsilon + o(\varepsilon),$$

where, as with a in (i), \tilde{a}_\pm takes different values corresponding to different roots.

By summing over *all* the roots, we find that

$$-\beta_1^{(1)} = \sum a + \sum (\tilde{a}_+ + \tilde{a}_-). \tag{57}$$

Moreover, substituting the expressions for $\lambda(\varepsilon)$ given in (i) and (ii) into (54) yields

$$2\delta|\lambda_0|\varepsilon \geq (2 \operatorname{Re} \bar{\lambda}_0 a)\varepsilon + o(\varepsilon) \tag{58}$$

in case (i) and

$$2\delta|\lambda_0|\varepsilon \geq (|r_h| + 2 \operatorname{Re} \bar{\lambda}_0 \tilde{a}_\pm)\varepsilon + o(\varepsilon) \tag{59}$$

in case (ii). Now summing these inequalities over *all* the roots gives the inequality

$$2\delta t_0 |\lambda_0| \varepsilon \geq \left[2 \sum_{h=1}^m |r_h| + 2 \operatorname{Re} \bar{\lambda}_0 (\sum a + \sum (\tilde{a}_+ + \tilde{a}_-)) \right] \varepsilon + o(\varepsilon), \tag{60}$$

where the sums without explicit indexing indicates summing over all the roots, while the factor 2 appearing in the front of $\sum_{h=1}^m |r_h|$ reflects the fact that there are two roots for each $h = 1, 2, \dots, m$. Now, as explained in lemma 1,

$$\sum_{h=1}^m |r_h| \geq \left| \sum_{h=1}^m r_h \right| = |\beta_2^{(1)}|. \tag{61}$$

Combining this inequality with (57) and (60) gives

$$\delta|\lambda_0| \geq \frac{1}{t_0} (|\beta_2^{(1)}| - \operatorname{Re} \bar{\lambda}_0 \beta_1^{(1)}) + \frac{o(\varepsilon)}{\varepsilon}. \tag{62}$$

Letting $\varepsilon \downarrow 0$, we obtain the inequality

$$\delta|\lambda_0| \geq \frac{1}{t_0} (|\beta_2^{(1)}| - \operatorname{Re} \bar{\lambda}_0 \beta_1^{(1)}). \tag{63}$$

We now consider the two cases $\rho(z^0) = 0$ and $\rho(z^0) \neq 0$ separately. If $\rho(z^0) = 0$, then inequality (63) implies that $\beta_2^{(1)} = 0$. Thus, $\rho^h(z^0; d) = +\infty$ unless $\beta_j^{(1)} = 0$ for $j = 2, \dots, t_0$. Furthermore, observe that

$$|\sum \lambda(\varepsilon)| \leq \sum |\lambda(\varepsilon)| < t_0 \delta \varepsilon,$$

for all $\varepsilon \in [0, \varepsilon_0]$ where the sum is taken over all branches $\lambda(\varepsilon)$. Letting $\delta \downarrow \rho^h(z^0; d)$ yields (48). On the other hand, if $\rho(z^0) \neq 0$, then (51) follows immediately from (63) by letting $\delta \downarrow \rho^h(z^0; d)$.

Next, suppose that the rank of $c'(z^0)$ is t_0 and that (49) and (50) hold (otherwise equality in either (48) or (51) is trivially satisfied). We will only consider the case $\rho(z^0) \neq 0$ since the case $\rho(z^0) = 0$ follows in a similar manner. Moreover, in the case $t_0 = 1$, we take $\beta_2^{(1)} = 0$ as usual. Consider the coefficients of the powers of $(\lambda - \lambda_0)$ in the polynomial

$$\begin{aligned} & ((\lambda - \lambda_0) + \sigma\varepsilon)^{t_0-2} [(\lambda - \lambda_0) + \sqrt{-\beta_2^{(1)}}\sqrt{\varepsilon} + \tau\varepsilon][(\lambda - \lambda_0) - \sqrt{-\beta_2^{(1)}}\sqrt{\varepsilon} + \tau\varepsilon] \\ &= (\lambda - \lambda_0)^{t_0} + \beta_1^{(1)}\zeta(\lambda - \lambda_0)^{t_0-1} + (\beta_2^{(1)}\zeta + O(\zeta^2))(\lambda - \lambda_0)^{t_0-2} + \dots, \end{aligned} \tag{64}$$

where σ and τ are defined by the expressions

$$\sigma = \frac{-1}{t_0} [|\beta_2^{(1)}| - \operatorname{Re} \bar{\lambda}_0 \beta_1^{(1)}] \frac{\lambda_0}{|\lambda_0|^2}$$

and

$$\tau = \frac{1}{2}[\beta_1^{(1)} - (t_0 - 2)\sigma].$$

These coefficients are chosen so that not only are (20) and (21) satisfied with $y_0 = \bar{\lambda}_0$, but also the expansion of each root satisfies

$$|\lambda(\varepsilon)|^2 = |\lambda_0|^2 + \frac{2}{t_0} [|\beta_2^{(1)}| - \operatorname{Re} \bar{\lambda}_0 \beta_1^{(1)}] \varepsilon + o(\varepsilon).$$

The proof now follows the argument given in the proof of theorem 2, that is, one uses the rank condition and the implicit function theorem to establish the existence of a curve $\gamma \in \Gamma(z^0, d)$ such that (13) has the same coefficients as (64). With this choice of γ equality holds in (51). ■

The main theorem of this section now follows from theorem 8. It is derived from theorem 8 in the same way that theorem 3 was derived from theorem 2 and so its proof is omitted.

THEOREM 9. Define ρ by (2) and (3) where P has the representation (10) near $z^0 \in \mathbb{C}^r$ with each μ_k given by (11). Choose $d \in \mathbb{C}^r$, and define

$$\mathfrak{R}_0(z^0) = \{\lambda_k \in \Sigma(z^0): |\lambda_k| = \rho(z^0)\}.$$

We consider the two cases $\rho(z^0) = 0$ and $\rho(z^0) \neq 0$ separately.

(1) Suppose $\rho(z^0) = 0$ so that $\mathfrak{R}_0(z^0) = \Sigma(z^0) = \{\lambda_1\}$ and $t_1 = n$. If any one of the conditions

$$c'_{1j}(z^0)d = 0, \quad j = 2, \dots, n,$$

is violated, then

$$\rho^h(z^0; d) = +\infty; \tag{65}$$

otherwise,

$$\rho^h(z^0; d) \geq \frac{1}{n} |c'_{11}(z^0)d|. \tag{66}$$

(2) Suppose $\rho(z^0) \neq 0$. If for some $\lambda_0 \in \mathfrak{R}_0(z^0)$ any one of the conditions

$$\operatorname{Re} \bar{\lambda}_0^2 c'_{k2}(z^0)d \geq 0, \quad \operatorname{Im} \bar{\lambda}_0^2 c'_{k2}(z^0)d = 0, \tag{67}$$

$$c'_{kj}(z^0)d = 0, \quad j = 3, \dots, t_k, \tag{68}$$

is violated, then

$$\rho^h(z^0; d) = +\infty; \tag{69}$$

otherwise,

$$\rho^h(z^0; d) \geq \max\{\Psi(\lambda_k): \lambda_k \in \mathfrak{R}_0(z^0)\}, \tag{70}$$

where

$$\Psi(\lambda_k) = \frac{1}{t_k \rho(z^0)} [|c'_{k2}(z^0)d| - \operatorname{Re} \bar{\lambda}_0 c'_{k1}(z^0)d]$$

with c_{k2} understood to be the zero map if $t_k = 1$.

Moreover, if the vectors $\{c'_{kj}(z^0): \lambda_k \in \mathcal{R}_1(z^0, d), j = 1, \dots, t_k\}$ are linearly independent, where $\mathcal{R}_1(z^0, d)$ is the set of $\lambda_k \in \mathcal{R}_0(z^0)$ such that $\Psi(\lambda_k) = \max\{\Psi(\xi): \xi \in \mathcal{R}_0(z^0)\}$ when $\rho(z^0) \neq 0$ and $\mathcal{R}_1(z^0, d) = \Sigma(z^0)$ otherwise, then equality holds in (66) or (70) depending on whether $\rho(z^0) = 0$ or $\rho(z^0) \neq 0$ holds, respectively.

One can now apply theorem 9 in conjunction with theorem 5 to obtain a result for $\rho^h(z^0; d)$, where ρ is defined by (2)–(4), in terms of matrix elements, using the Jordan structure described in Section 3. We omit the details, but we briefly consider the case where $\mathcal{R}_0(z^0)$ consists of the single nonzero eigenvalue λ_1 , with multiplicity t_1 . Let $C = \sum_{q=1}^{\nu} d_q B_{q1}$. Then the necessary condition for $\rho^h(z^0; d)$ to be finite reduces to

$$\operatorname{Re} \bar{\lambda}_1 \operatorname{tr}^{(2)} C \leq 0, \quad \operatorname{Im} \bar{\lambda}_1 \operatorname{tr}^{(2)} C = 0, \quad \operatorname{tr}^{(j)} C = 0, \quad j = 3, \dots, t_1$$

and, if this condition holds,

$$\rho^h(z^0; d) \geq \frac{1}{t_1 |\lambda_1|} [\operatorname{Re} \bar{\lambda}_1 \operatorname{tr}^{(1)} C - |\operatorname{tr}^{(2)} C|] \tag{71}$$

with equality if the linear independence condition also holds.

Note that, as before, the linear independence condition *cannot* hold if any of the eigenvalues λ_k in $\mathcal{R}_1(z^0, d)$ is derogatory. Thus, in this case, one should not expect the lower bound for the directional derivative to be achieved. Indeed, when $\mathcal{R}_0(z^0)$ contains only semisimple eigenvalues, we obtain the following result.

THEOREM 10. Define ρ by (2)–(4). If $\mathcal{R}_0(z^0)$ contains only semisimple eigenvalues, then for every $d \in \mathbb{C}^{\nu}$ the ordinary directional derivative $\rho'(z_0; d)$ exists, equals $\rho^h(z_0; d)$, and satisfies

$$\rho'(z^0; d) = \begin{cases} \max_{1 \leq l \leq n_1} |\lambda'_{1l}| & \text{if } \rho(z_0) = 0, \\ \frac{1}{\rho(z_0)} \max_{k \in \mathcal{R}_0(z^0)} \max_{1 \leq l \leq n_k} \operatorname{Re} \bar{\lambda}_k \lambda'_{kl} & \text{if } \rho(z_0) \neq 0, \end{cases} \tag{72}$$

where $\lambda'_{kl}, l = 1, \dots, t_k$ are the eigenvalues of $\sum_{q=1}^{\nu} d_q B_{qk}$ and the matrices B_{qk} are defined in theorem 5.

Proof. It follows from [6, Section II.2.3; 8, Section 3] in the same way as the proof of theorem 7. ■

Therefore, if, for example, $\rho(z^0) \neq 0$, $\mathcal{R}_0(z^0)$ contains only the single element λ_1 , and λ_1 is semisimple, then the lower bound on $\rho^h(z_0; d)$ given by (71) is the *average* of the values $(1/|\lambda_1|) \operatorname{Re} \bar{\lambda}_1 \lambda'_{1l}, l = 1, \dots, t_1$, whereas $\rho^h(z^0; d)$ is the *maximum* of these quantities.

5. CONCLUDING REMARKS

We have defined a new directional derivative, based on analytic curves tangent to the direction, which is suitable for the analysis of the spectral abscissa and spectral radius functions, and we have applied a very classical technique, namely the Puiseux–Newton diagram, to obtain necessary conditions for the directional derivative to be finite and, in the

case where these hold, lower bounds for its value. Moreover, as a consequence of the fact that these lower bounds were derived from properties of the characteristic polynomial, it was observed that, subject to a nondegeneracy condition, the lower bounds are sharp when the eigenvalues corresponding to the value of either α or ρ are nonderogatory. This is the most interesting case since nonderogatory eigenvalues are the most generic. These results nicely complement those obtained by Overton and Womersley [8] for the semisimple case, since, when multiple eigenvalues occur, the nonderogatory and semisimple cases lie at opposite structural extremes. In the semisimple case, the new directional derivative reduces to the ordinary directional derivative, and its value is the *maximum* of certain quantities depending on the derivative of $A(z)$. In the nonderogatory case the directional derivative may be infinite, but in the case where the necessary conditions for it to be finite hold, the lower bound on its value is the maximum of the *average* of subsets of the *same* quantities, in the case of the spectral abscissa, and closely related quantities, in the case of the spectral radius.

REFERENCES

1. LYAPUNOV A. M. (LIAPOUNOFF), Le problème général de la stabilité du mouvement, *Annals of Mathematics Studies*, No. 17. Princeton University Press, Princeton (1949).
2. BOYD S. & YANG Q., Structured and simultaneous Lyapunov functions for system stability problems, *Int. J. Control* **49**, 2215–2240 (1989).
3. HERSHKOWITZ D., Recent directions in matrix stability, *Linear Algebra Appl.* **171**, 161–186 (1992).
4. MILLER K. S., *Linear Difference Equations*. Benjamin, New York (1968).
5. GANTMACHER F. R., *Matrix Theory*, Vol. II. Chelsea, New York (1959).
6. KATO T., *A Short Introduction to Perturbation Theory for Linear Operators*. Springer, New York (1982).
7. DEMYANOV V. F. & MALOZEMOV V. N., *Introduction to Minimax*. Wiley, New York (1974).
8. OVERTON M. L. & WOMERSLEY R. S., On minimizing the spectral radius of a nonsymmetric matrix function: optimality conditions and duality theory, *SIAM J. Matrix Anal. Appl.* **9**, 473–498 (1988).
9. ARNOLD V. I., On matrices depending on parameters, *Russ. math. Survs* **26**(2), 29–43 (1971).
10. NEWTON I., *Enumeration of Lines of Third Order/ Generation of Curves by Shadows/ Organic Description of Curves, and Construction of Equations/* Translation from Latin by C. R. M. Talbot. Bohn, London (1860).
11. VAINBERG M. M. & TRENTOGEN V. A., *Theory of Branching of Solutions of Non-linear Equations*. Noordhoff, Leyden (1974).
12. CLARKE F. H., *Optimization and Nonsmooth Analysis*. Wiley, New York (1983).
13. IOFFE A. D., Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps, *Nonlinear Analysis* **8**, 517–539 (1984).
14. MICHEL P. & PENOT J.-P., Calcul sous-différentiel pour des fonctions lipschitziennes et pour non lipschitziennes, *C.r. Acad. Sci. Paris* **298**, 269–272 (1984).
15. MORDUKHOVICH B. S., *Approximation Methods in Optimization and Control Problems*. Nauka, Moscow (1988). (In Russian.)
16. ROCKAFELLAR R. T., Extensions of subgradient calculus with applications to optimization, *Nonlinear Analysis* **9**(7), 665–698 (1985).
17. BURKE J. V. & OVERTON M. L., On the subdifferentiability of functions of a matrix spectrum II: subdifferential formulas in *Nonsmooth Optimization: Methods and Applications* (Edited by F. GIANNESI), pp. 19–29. Proceedings of a conference held in Erice, Italy, June 1991, Gordon & Breach, Philadelphia (1992).
18. CLARKE F. H., Necessary conditions for nonsmooth problems in optimal control and the calculus of variations, Ph.D. thesis, University of Washington, Seattle (1973).
19. BAUMGÄRTEL H., *Analytic Perturbation Theory for Matrices and Operators*. Birkhäuser, Boston (1985).
20. PUISEUX V., Recherches sur les fonctions algébriques, *J. Math. pures appl.* **51**, 465–581 (1850).
21. LIDSKII V. B., On the perturbation theory for nonselfadjoint operators, *Zvycisl. Mat. i Mat. Fiz.* **6**, 52–60 (1966). (In Russian.)
22. LANGER H. & NAJMAN B., Remarks on the perturbation of analytic matrix functions III, *Integr. Eqns Op. Theor.* **15**, 796–806 (1992).
23. GOHBERG I., LANCASTER P. & RODMAN L., *Matrix Polynomials*. Academic Press, New York (1982).
24. BURKE J. V. & OVERTON M. L., Stable perturbations of nonsymmetric matrices, *Linear Algebra Appl.* **171**, 249–273 (1992).
25. LEVANTOVSKII L. V., The boundary of a set of stable matrices, *Russ. math. Survs* **35**, 249–250 (1980).

26. HÖRMANDER L., *An Introduction to Complex Analysis in Several Variables*. North-Holland, New York (1990).
27. BURKE J. V. & OVERTON M. L., On the subdifferentiability of functions of a matrix spectrum, I: mathematical foundations, in *Nonsmooth Optimization: Methods and Applications* (Edited by F. GIANNESI), pp. 11-18. Proceedings of a conference held in Erice, Italy, June 1991. Gordon & Breach, Philadelphia (1992).
28. GOLUB G. H. & VAN LOAN C., *Matrix Computations*, 2nd Edition. Johns Hopkins University Press, Baltimore (1989).