

Variational Analysis of the Spectral Abscissa at a Matrix with a Nongeneric Multiple Eigenvalue

Sara Grundel · Michael L. Overton

the date of receipt and acceptance should be inserted later

Abstract The spectral abscissa is a fundamental map from the set of complex matrices to the real numbers. Denoted α and defined as the maximum of the real parts of the eigenvalues of a matrix X , it has many applications in stability analysis of dynamical systems. The function α is nonconvex and is non-Lipschitz near matrices with multiple eigenvalues. Variational analysis of this function was presented in [BO01], including a complete characterization of its regular subgradients and necessary conditions which must be satisfied by all its subgradients. A complete characterization of all subgradients of α at a matrix X was also given for the case that all active eigenvalues of X (those whose real part equals $\alpha(X)$) are nonderogatory (their geometric multiplicity is one) and also for the case that they are all nondefective (their geometric multiplicity equals their algebraic multiplicity). However, necessary and sufficient conditions for all subgradients in all cases remain unknown. In this paper

Supported in part by National Science Foundation Grant DMS-0714321 at Courant Institute of Mathematical Sciences, New York University

Supported in part by National Science Foundation Grant DMS-1016325

Sara Grundel
Computational Methods in Systems and Control Theory,
Max Planck Institute for Dynamics of Complex Technical Systems,
Sandtorstr. 1,
39106 Magdeburg
Tel.: +49-391-6110-805
Fax: +49-391-6110-453
E-mail: grundel@mpi-magdeburg.mpg.de

Michael L. Overton
Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, NY, 10012
E-mail: overton@cs.nyu.edu

we present necessary and sufficient conditions for the simplest example of a matrix X with a derogatory, defective multiple eigenvalue.

1 Introduction

Let $\mathbb{C}^{n \times n}$ denote the Euclidean space of $n \times n$ complex matrices. Define the *spectral abscissa* $\alpha : \mathbb{C}^{n \times n} \rightarrow \mathbb{R}$ by

$$\alpha(X) = \max \{ \operatorname{Re} \lambda : \lambda \in \sigma(X) \}$$

where Re denotes real part and $\sigma(X)$ denotes the spectrum, or set of eigenvalues, of X . The spectral abscissa, along with other related spectral functions such as the spectral radius, is a fundamental concept with many applications in stability analysis of dynamical systems. Spectral functions are nonconvex and non-Lipschitz near matrices with multiple eigenvalues. Variational analysis for a broad class of spectral functions was presented in [BO01]. In this paper we present some extensions of these results for the spectral abscissa.

Define an eigenvalue $\lambda \in \sigma(X)$ to be *active* if $\operatorname{Re} \lambda = \alpha(X)$. The algebraic multiplicity of λ is its multiplicity as a root of the characteristic polynomial $\det(X - zI)$ while its geometric multiplicity is the number of linearly independent eigenvectors v satisfying $Xv = \lambda v$. The latter is always less than or equal to the former. When the algebraic multiplicity is one, λ is said to be *simple* and when the geometric multiplicity is one λ is said to be *nonderogatory*. If the algebraic multiplicity equals the geometric multiplicity, λ is said to be *nondefective* or *semisimple*. Simple eigenvalues are nonderogatory and nondefective. A nonderogatory eigenvalue λ with multiplicity m corresponds to a single $m \times m$ Jordan block in the Jordan normal form of X , while a nondefective eigenvalue λ with multiplicity m corresponds to m scalar blocks in the Jordan form of X . The set of matrices with a given Jordan block structure defines a submanifold of $\mathbb{C}^{n \times n}$ whose properties are well known [Arn71]. Given that an eigenvalue λ has multiplicity m , the most generic Jordan structure is a single block, meaning λ is nonderogatory.

The analysis in [BO01] included a complete characterization of all subgradients of the spectral abscissa α at a matrix X for the case that all active eigenvalues of X are nonderogatory, showing also that this is exactly the case when α is regular at X . A complete characterization was also given for the case that all active eigenvalues are nondefective. However, necessary and sufficient conditions for all subgradients in the case of active eigenvalues that are both derogatory and defective remain unknown. In this paper we present necessary and sufficient conditions for the simplest example of a matrix X with a derogatory, defective active eigenvalue.

The paper is organized as follows. In Section 2 we review some of the results of Burke and Overton [BO01] for variational analysis of the spectral abscissa. We then state our main result for the simplest derogatory, defective example in Section 3. The proof is given in Sections 4 and 5. We discuss the horizon subgradients in Section 6. We make some concluding remarks in Section 7.

2 Review of Known Results

We start by defining subgradients in the sense of Mordukhovich [Mor76] as expounded in [RW98, Chapter 10]. We treat $\mathbb{C}^{n \times n}$ as a Euclidean space with the real inner product

$$\langle Y, Z \rangle = \operatorname{Re} \operatorname{tr} Y^* Z, \quad (1)$$

where $*$ denotes the complex conjugate transpose operation and tr denotes trace. Since α is a continuous function, its subgradients are defined as follows. A matrix $Y \in \mathbb{C}^{n \times n}$ is a *regular subgradient* of α at $X \in \mathbb{C}^{n \times n}$ (written $Y \in \hat{\partial}\alpha(X)$) if

$$\liminf_{Z \rightarrow 0} \frac{\alpha(X + Z) - \alpha(X) - \langle Y, Z \rangle}{\|Z\|} \geq 0.$$

A matrix $Y \in \mathbb{C}^{n \times n}$ is a *subgradient* of α at X (written $Y \in \partial\alpha(X)$) if there exist sequences X_i and Y_i in $\mathbb{C}^{n \times n}$ satisfying

$$X_i \rightarrow X, \quad Y_i \in \hat{\partial}\alpha(X_i), \quad Y_i \rightarrow Y. \quad (2)$$

A matrix $Y \in \mathbb{C}^{n \times n}$ is a *horizon subgradient* of α at X (written $Y \in \partial^\infty\alpha(X)$) if $Y = 0$ or (2) holds with the last condition replaced by $s_i Y_i \rightarrow Y$, where s_i is a positive real sequence converging to zero.

Let us establish notation for the Jordan normal form. A nonsingular matrix $P \in \mathbb{C}^{n \times n}$ transforms $X \in \mathbb{C}^{n \times n}$ to *Jordan form* if

$$P^{-1}XP = J = \begin{pmatrix} J^{(1)} & & \\ & \ddots & \\ & & J^{(p)} \end{pmatrix}, \text{ where } J^{(j)} = \begin{pmatrix} J_1^{(j)} & & \\ & \ddots & \\ & & J_{q^{(j)}}^{(j)} \end{pmatrix} \quad (3)$$

$$\text{with } J_k^{(j)} = \begin{pmatrix} \mu_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \mu_j \end{pmatrix}, \quad k = 1, \dots, q^{(j)}, \quad j = 1, \dots, p. \quad (4)$$

Here μ_1, \dots, μ_p denote the distinct eigenvalues of X . Each μ_j corresponds to $q^{(j)}$ Jordan blocks $J_k^{(j)}$ with size $m_k^{(j)} \times m_k^{(j)}$. Note that $q^{(j)}$ is the geometric multiplicity of μ_j , while

$$m^{(j)} = \sum_{k=1}^{q^{(j)}} m_k^{(j)}$$

is its algebraic multiplicity. The size of the largest Jordan block for μ_j is denoted

$$n^{(j)} = \max_{k=1, \dots, q^{(j)}} m_k^{(j)}.$$

The eigenvalue μ_j is said to be *nonderogatory* if $q^{(j)} = 1$ and *nondefective* if $n^{(j)} = 1$ (and hence $q^{(j)} = m^{(j)}$). Finally, let

$$\mathcal{A} = \{j \mid \operatorname{Re} \mu_j = \alpha(X)\}.$$

The eigenvalue μ_j is said to be *active* if $j \in \mathcal{A}$.

We now review several results from [BO01] that we will use. Since we need these only for the spectral abscissa, we present them more succinctly than they appear in [BO01]. We start with necessary conditions for subgradients and horizon subgradients. In what follows, the structure in W imposed by the structure in the Jordan form J arises from the commutativity condition $XY^* = Y^*X$ derived in [BO01, Theorem 2.1].

Theorem 1 [BO01, Corollary 8.1] *If $Y \in \partial\alpha(X)$ or $Y \in \partial^\infty\alpha(X)$, then any P satisfying (3), (4) also satisfies*

$$P^*YP^{-*} = W = \begin{pmatrix} W^{(1)} & & \\ & \ddots & \\ & & W^{(p)} \end{pmatrix}, \quad (5)$$

with $W^{(j)} = 0$ if $j \notin \mathcal{A}$, and otherwise

$$W^{(j)} = \begin{pmatrix} W_{11}^{(j)} & \cdots & W_{1q^{(j)}}^{(j)} \\ \vdots & \ddots & \vdots \\ W_{q^{(j)}1}^{(j)} & \cdots & W_{q^{(j)}q^{(j)}}^{(j)} \end{pmatrix}, \quad (6)$$

where $W_{rs}^{(j)}$ is a rectangular $m_r^{(j)} \times m_s^{(j)}$ lower triangular Toeplitz matrix, $r = 1, \dots, q^{(j)}$, $s = 1, \dots, q^{(j)}$, $j = 1, \dots, p$. By this we mean that the value of the k, ℓ entry in each $W_{rs}^{(j)}$ depends only on the difference $k - \ell$ (is constant along the diagonals), and is zero if $k < \ell$ or $m_r^{(j)} - k > m_s^{(j)} - \ell$ (is zero above the main diagonal, drawn either from the top left of the block, or from the bottom right). Finally, if $Y \in \partial\alpha(X)$ then the eigenvalues of Y (equivalently of W) are real, nonnegative, and sum to one, while if $Y \in \partial^\infty\alpha(X)$, it must be nilpotent, meaning that its eigenvalues (equivalently the eigenvalues of W) are all zero.

See [Arn71, Sec. 4.2] or [OW88] for illustrations of the block structure in (6).

In the case of *regular* subgradients, necessary and sufficient conditions are known:

Theorem 2 [BO01, Theorem 7.2] *If $Y \in \hat{\partial}\alpha(X)$ then the diagonal blocks of $W = P^*YP^{-*}$ in (5) satisfy $W^{(j)} = 0$ if $j \notin \mathcal{A}$ and otherwise*

$$W^{(j)} = \begin{pmatrix} W_{11}^{(j)} & & \\ & \ddots & \\ & & W_{q^{(j)}q^{(j)}}^{(j)} \end{pmatrix}$$

with

$$W_{kk}^{(j)} = \begin{pmatrix} \theta_1^{(j)} & & & \\ \theta_2^{(j)} & \cdot & & \\ \cdot & \cdot\cdot & & \\ \cdot & \cdot\cdot & \cdot & \\ \theta_{m_k}^{(j)} & \cdot\cdot & \theta_2^{(j)} & \theta_1^{(j)} \end{pmatrix}, k = 1, \dots, q^{(j)}, \quad j = 1, \dots, p,$$

for some $\theta_\ell^{(j)}$, $\ell = 1, \dots, n^{(j)}$. Thus, for each $j \in \mathcal{A}$, the block $W^{(j)}$ is itself block diagonal with square lower triangular Toeplitz blocks, with the entries on the diagonals of the Toeplitz blocks constant not only within each block, but also across all $q^{(j)}$ blocks. Furthermore, in accordance with the condition on the eigenvalues of Y , we have

$$\theta_1^{(j)} \in \mathbb{R}, \quad \theta_1^{(j)} \geq 0, \quad \sum_{j \in \mathcal{A}} m^{(j)} \theta_1^{(j)} = 1$$

and also the additional condition

$$\operatorname{Re} \theta_2^{(j)} \geq 0, \quad j \in \mathcal{A}.$$

Finally, only those Y satisfying the conditions given above lie in $\hat{\partial}\alpha(X)$.

Necessary and sufficient conditions for Y to be a subgradient are known for the case that all active eigenvalues of X are nonderogatory. In fact, this is exactly the case that α is *regular* at X : all its subgradients are regular, and the horizon subgradients satisfy a recession cone condition.

Theorem 3 [BO01, Theorem 8.2] *Suppose that $Y \in \partial\alpha(X)$ and that all active eigenvalues of X are nonderogatory, so that the corresponding matrices $J^{(j)}$ are all full Jordan blocks of order $m^{(j)}$, and hence the diagonal blocks of $W = P^*YP^{-*}$ in (5) are all lower triangular Toeplitz, i.e.,*

$$W^{(j)} = \begin{pmatrix} \theta_1^{(j)} & & & \\ \theta_2^{(j)} & \cdot & & \\ \cdot & \cdot\cdot & & \\ \cdot & \cdot\cdot & \cdot & \\ \theta_{m^{(j)}}^{(j)} & \cdot\cdot & \theta_2^{(j)} & \theta_1^{(j)} \end{pmatrix},$$

for some $\theta_\ell^{(j)}$, $\ell = 1, \dots, m^{(j)}$. We have $W^{(j)} = 0$ if $j \notin \mathcal{A}$, and in accordance with the condition on the eigenvalues of Y we have

$$\theta_1^{(j)} \in \mathbb{R}, \quad \theta_1^{(j)} \geq 0, \quad \sum_{j \in \mathcal{A}} m^{(j)} \theta_1^{(j)} = 1$$

as well as the additional condition

$$\operatorname{Re} \theta_2^{(j)} \geq 0, \quad j \in \mathcal{A}.$$

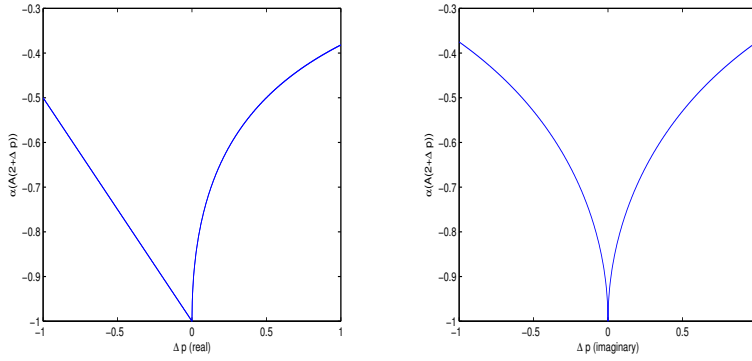


Fig. 1 The spectral abscissa of the damped linear oscillator. Left: $\alpha(A(2 + \Delta p))$ for real Δp . Right: same, for imaginary Δp .

Only those Y satisfying the conditions given above are subgradients of α at X , and all these subgradients are regular, so $\partial\alpha(X) = \partial\alpha(X)$. Finally $\partial^\infty\alpha(X)$, the set of horizon subgradients of α at X , is the recession cone of $\partial\alpha(X)$, that is, horizon subgradients satisfy the same conditions as subgradients except that $\theta_1^{(j)} = 0$, $j = 1, \dots, p$.

Necessary and sufficient conditions for Y to be a subgradient are also known for the case that all active eigenvalues of X are nondefective.

Theorem 4 [BO01, Theorem 8.3] *Suppose that $Y \in \partial\alpha(X)$ and that all active eigenvalues of X are nondefective, so that the corresponding $J^{(j)} = \mu_j I$, and hence the diagonal blocks of $W = P^* Y P^{-*}$ in (5) have no particular structure. We have $W^{(j)} = 0$ if $j \notin \mathcal{A}$, and the eigenvalues of Y (equivalently of W) are all real, nonnegative, and sum to one. Only those matrices satisfying these conditions lie in $\partial\alpha(X)$. Furthermore, the set of horizon subgradients of α at X consists of the matrices Y satisfying the same conditions, except that the eigenvalues of Y (equivalently of W) are all zero.*

We close this section by mentioning some examples of the usefulness of this theory. In applications, often one wants to minimize a spectral function over a family of matrices $A(p)$ where p is a vector of parameters. A simple but instructive example is maximizing the asymptotic decay rate for the damped linear oscillator $u'' + pu' + u$ over a single damping parameter p [BLO01]. Although for physical reasons one would normally assume p is real, for technical reasons related to the chain rule used below we take p to be a complex parameter. This yields the following spectral abscissa minimization problem:

$$\min_{p \in \mathbb{C}} \alpha(A(p)) \text{ with } A(p) = \begin{pmatrix} 0 & 1 \\ -1 & -p \end{pmatrix}.$$

Figure 1 plots $\alpha(A(p))$ for $p = 2 + \Delta p$ with Δp in the real interval $[-1, 1]$ in the left panel and Δp in the imaginary interval $[-i, i]$ in the right panel. The

minimizer is $\tilde{p} = 2$, for which $A(\tilde{p})$ has the Jordan form

$$P^{-1}A(\tilde{p})P = J = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \text{ with } P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

The only active eigenvalue of $A(\tilde{p})$ is the double, nonderogatory eigenvalue -1 . According to Theorem 1, the subgradients of α at $A(\tilde{p})$ are given by

$$\partial\alpha(X) = \left\{ P^{-*} \begin{pmatrix} \frac{1}{2} & 0 \\ \theta & \frac{1}{2} \end{pmatrix} P^* = \begin{pmatrix} \frac{1}{2} + \theta & -\theta \\ \theta & \frac{1}{2} - \theta \end{pmatrix} \text{ with } \operatorname{Re} \theta \geq 0 \right\}$$

while the horizon subgradients are given by

$$\partial^\infty\alpha(X) = \left\{ P^{-*} \begin{pmatrix} 0 & 0 \\ \theta & 0 \end{pmatrix} P^* = \begin{pmatrix} \theta & -\theta \\ \theta & -\theta \end{pmatrix} \text{ with } \operatorname{Re} \theta \geq 0 \right\}.$$

Furthermore, α is regular at $A(\tilde{p})$. Noting that the only horizon subgradient orthogonal to the Jacobian $\nabla_p A$ in the inner product (1) is zero, we can apply a chain rule from nonsmooth analysis [RW98, Theorem 10.6] to conclude that

$$\begin{aligned} \partial(\alpha \circ A)(\tilde{p}) &= \{ \langle \nabla_p A, Y \rangle \text{ with } Y \in \partial\alpha(A(\tilde{p})) \} \\ &= \{ \theta - \frac{1}{2} \text{ with } \operatorname{Re} \theta \geq 0 \}. \end{aligned} \quad (7)$$

Projecting this set onto the real and imaginary axes we obtain $[-\frac{1}{2}, \infty)$ and $(-\infty, \infty)i$ respectively, as is consistent with the left and right panels of Figure 1. Note that the unbounded derivatives for real $p > \tilde{p}$ indicate a splitting of the double eigenvalue -1 into two real eigenvalues, one of which grows as $O(p - \tilde{p})^{1/2}$, while for real $p < \tilde{p}$, the double eigenvalue splits into a complex conjugate pair for which only the imaginary part grows as $O(\tilde{p} - p)^{1/2}$. This distinction is captured in the set of subgradients and horizon subgradients by the condition $\operatorname{Re} \theta \geq 0$. If we were to change the example to one with a nonderogatory eigenvalue with multiplicity three, we would see non-Lipschitz growth in every perturbation direction. Finally, since 0 is in the interior of $\partial(\alpha \circ A)(\tilde{p})$, we conclude that \tilde{p} is in fact a sharp local minimizer [BLO01, Prop. 4.3] of $\alpha \circ A$, as is indicated by Figure 1.

In the example just discussed the matrix family depends on only one parameter. For another simple example with multiple parameters, see [BLO01, Section 2]. For more interesting examples motivated by some long-standing open questions in control, see [BHLO06, Section III] and [HO06]. For all of these examples, an optimality analysis of the sort just described for the damped linear oscillator leads to the conclusion that a certain choice of parameters \tilde{p} gives a sharp local minimizer of the spectral abscissa of a specific matrix family.

Spectral radius optimization problems are also important in applications. For an example of a matrix family $A(p)$ describing the behavior of a Markov chain, see [GO07]. In this example, the largest eigenvalue in modulus is fixed at one and the goal is to minimize the largest of the complex moduli of the remaining eigenvalues in order to obtain the Markov chain with fastest asymptotic convergence. A local optimality analysis is done at a candidate optimal

matrix with many distinct active eigenvalues, one of them a double real non-derogatory eigenvalue and the others simple eigenvalues occurring in complex conjugate pairs, but all having the same modulus.

Although nonderogatory eigenvalues are the most generic, structure present in the matrix family may lead to local optimizers \tilde{p} for which $A(\tilde{p})$ has active eigenvalues which are both derogatory and defective. Two spectral radius optimization problems of this sort, arising in surface subdivision schemes with applications in computer graphics, are studied in [Gru11, Chapters 2 and 4]. In both problems several of the largest eigenvalues of a matrix family $A(p)$ are fixed and the largest of the moduli of the remaining eigenvalues is to be minimized. In one of these examples, the optimal matrix $A(\tilde{p})$ has one active eigenvalue associated with three Jordan blocks respectively having order 2, 1 and 1. In the second example, $A(\tilde{p})$ apparently has four Jordan blocks respectively having order 5, 3, 2 and 2. Both of these optimizers were found numerically; in the first case, the Jordan structure was then verified analytically but in the second case this was not possible given the complexity of the problem. In both these examples the active eigenvalue is apparently zero (definitely in the first case) so optimality conditions are not needed as no eigenvalue can have smaller modulus than zero. However, similar problems can be constructed for which analysis of optimality conditions is the only path to verifying local optimality.

Returning to the spectral abscissa α , in the case of derogatory active eigenvalues one cannot expect to use a chain rule as strong as the one exploited in (7), because α is not regular at matrices with a derogatory active eigenvalue. However, an inclusion of the form

$$\partial(\alpha \circ A)(\tilde{p}) \subset \{\langle \nabla_p A, Y \rangle \text{ for } Y \in \partial\alpha(A(\tilde{p}))\}$$

still applies [RW98, Theorem 10.6]. Hence the importance of obtaining necessary conditions for $Y \in \partial\alpha(A)$ in the derogatory, defective case that are as strong as possible. This provides motivation for the analysis in the remainder of the paper.

3 Main Result

When $n = 2$ the eigenvalues of X are either nonderogatory or nondefective, so the simplest example of a matrix with an eigenvalue that is both derogatory and defective occurs when $n = 3$, namely

$$X = J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (8)$$

This matrix is in Jordan form, so P in (3) is the identity matrix, and its only eigenvalue is $\mu_1 = 0$, with $q^{(1)} = 2$ Jordan blocks of sizes $m_1^{(1)} = 2, m_2^{(1)} = 1$.

According to Theorem 1, necessary conditions for Y to be a subgradient of α at J are that it satisfies

$$Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & e \end{pmatrix}$$

and that the eigenvalues of Y are real, nonnegative and sum to one. Exchanging the second and third rows and the second and third columns transforms Y to triangular form, so the eigenvalues of Y are its diagonal entries a , a , and e . Hence the necessary conditions for Y to be a subgradient reduce to

$$Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & 1 - 2a \end{pmatrix} \text{ with } a \in \left[0, \frac{1}{2}\right]. \quad (9)$$

Before continuing with this example, we note that this observation applies more generally, as stated in the following lemma.

Lemma 1 *Let Y be a subgradient of α at X and assume that no active eigenvalue of X has multiple Jordan blocks with the same size, that is, for each active eigenvalue μ_j , the block sizes $m_1^{(j)}, \dots, m_{q^{(j)}}^{(j)}$ are distinct. Then each of the diagonal blocks of $W = P^*YP^{-*}$ in (5) has its eigenvalues on its diagonal, and hence the diagonal entries of W must be real, nonnegative and sum to one.*

Proof. Since W is block diagonal with zero blocks corresponding to inactive eigenvalues, we need only prove the result for each block $W^{(j)}$ corresponding to an active eigenvalue. Fix j , set $V = W^{(j)}$ and assume without loss of generality that $m_1^{(j)} > m_2^{(j)} > \dots > m_{q^{(j)}}^{(j)}$. We will show that there exists a permutation σ , that is a bijection from $\{1, \dots, m^{(j)}\}$ to $\{1, \dots, m^{(j)}\}$, with the property that the permuted block U , defined by $U_{i,i'} = V_{\sigma(i)\sigma(i')}$, is lower triangular. Since U and V have the same diagonal entries (in a different order), this will show that the eigenvalues of V are on its diagonal.

The structure of $V = W^{(j)}$ is shown in (6). Define the map

$$(b, k) : \{1, \dots, m^{(j)}\} \rightarrow \{1, \dots, q^{(j)}\} \times \{1, \dots, m_1^{(j)}\}$$

as follows: $b(i)$ is the block number corresponding to row and column i of V and $k(i)$ is the corresponding index within the block. Thus, for example, for all i we have $m_1^{(j)} + \dots + m_{b(i)-1}^{(j)} + k(i) = i$. We know from the comments following (6) that $V_{i,i'} = 0$ if either $k(i) < k(i')$ or $m_{b(i)}^{(j)} - k(i) > m_{b(i')}^{(j)} - k(i')$. So, we choose σ so that $\sigma(i) < \sigma(i')$ when $k(i) < k(i')$ or when $k(i) = k(i')$ and $m_{b(i)}^{(j)} > m_{b(i')}^{(j)}$. This last inequality is equivalent to $b(i) < b(i')$ and hence to $i < i'$. For example, if the block sizes are 4, 2 and 1, then, for $i = 1, \dots, 7$, $b(i)$ takes the values 1,1,1,1,2,2,3, $k(i)$ takes the values 1,2,3,4,1,2,1, and $\sigma(i)$ takes the values 1,5,7,2,6,3,4. This ensures that the permuted block U is lower triangular.

Since this argument holds for all active eigenvalues μ_j , we have shown that all the eigenvalues of W are on its diagonal. That the diagonal entries must be real, nonnegative and sum to one then follows from the last part of Theorem 1. \blacksquare

Returning to the example J in (8), first note that Theorem 2 says that necessary and sufficient conditions for Y to be a *regular* subgradient of α at J are

$$Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ w & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \text{ with } \operatorname{Re} w \geq 0. \quad (10)$$

We now present necessary and sufficient conditions for Y to be a subgradient of α at J :

Theorem 5 Define $g_{a,b,c,d} = \operatorname{Re} \left(b - \frac{cd}{1-3a} \right)$ and with it the set

$$\gamma(J) = \left\{ \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & 1-2a \end{pmatrix} \text{ with } a \in [0, \frac{1}{2}] \text{ and, if } a \neq \frac{1}{3}, g_{a,b,c,d} \geq 0 \right\} \quad (11)$$

This set is precisely the set of subgradients of α at J , i.e.,

$$\partial\alpha(J) = \gamma(J).$$

We break the proof into two parts, first showing that $\partial\alpha(J) \subset \gamma(J)$ in the next section and then showing the reverse inclusion in the following section.

4 Proof of the Inclusion $\partial\alpha(J) \subset \gamma(J)$

The method we use to prove Theorem 5 is a very direct one. We know that a subgradient Y must have the form (9). Since subgradients are limits of regular subgradients we study sequences $X_i \rightarrow J$ and $Y_i \rightarrow Y$ with $Y_i \in \hat{\partial}\alpha(X_i)$. The set of all such Y is the set of subgradients. There are only finitely many possible Jordan structures for 3×3 matrices, so we assume w.l.o.g. that in each sequence all X_i have the same Jordan structure (otherwise, we can consider a subsequence). We will go through each of the possible Jordan structures and discuss what limits are possible for Y_i , thereby establishing the new necessary conditions for Y given in the theorem statement. We will then prove that these conditions are sufficient in Section 5.

In dimension 3 there are 9 possible Jordan structures for $J_i = P_i^{-1}X_iP_i$, assuming in what follows that $\operatorname{Re} \alpha_i \geq \operatorname{Re} \beta_i \geq \operatorname{Re} \gamma_i$, namely

$$\begin{aligned} J_i^1 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \alpha_i \end{pmatrix} & J_i^2 &= \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \alpha_i \end{pmatrix} \\ J_i^3 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 1 \\ 0 & 0 & \alpha_i \end{pmatrix} & J_i^4 &= \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 1 \\ 0 & 0 & \alpha_i \end{pmatrix} \\ J_i^5 &= \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} & J_i^6 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} \\ J_i^7 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 1 \\ 0 & 0 & \beta_i \end{pmatrix} & J_i^8 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} \\ & & J_i^9 &= \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 0 \\ 0 & 0 & \gamma_i \end{pmatrix} \end{aligned}$$

Let us denote by $S_j(J)$ the set of all possible subgradient limits when X_i has Jordan form J_i^j . In the following we will go through each of the 9 cases and find necessary conditions for $Y \in S_j(J)$, $j = 1, \dots, 9$. For each j in turn, we write $X_i = P_i J_i^j P_i^{-1} \rightarrow J$ (note that X_i , Y_i and P_i also depend on j but we will suppress this dependence). Note that, in general, P_i^{-1} and P_i are not both bounded, but the eigenvalues of X_i , namely α_i, β_i and γ_i , must converge to 0. We will repeatedly exploit Theorem 2 which characterizes the structure of the regular subgradients $Y_i \in \partial\alpha(X_i)$. Note that each distinct Jordan structure J_i^j imposes a different structure on $W = P_i^* Y_i P_i^{-*}$.

Case 1

When the Jordan type of the sequence is J_i^1 we have $X_i = P_i(\alpha_i I)P_i^{-1} = \alpha_i I$, so it cannot converge to J and hence

$$S_1(J) = \emptyset.$$

Case 2

In this case we only get regular subgradients. Let

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \alpha_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ w_i & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} P_i^*$$

with $\operatorname{Re} w_i \geq 0$. Note that the given structure for Y_i is dictated by Theorem 2 since $Y_i \in \hat{\partial}\alpha(X_i)$. Now we do the following analysis:

$$\begin{aligned} Y_i^* &= P_i \begin{pmatrix} \frac{1}{3} & w_i^* & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} P_i^{-1} \\ &= P_i \left(\frac{1}{3}I + \begin{pmatrix} 0 & w_i^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right) P_i^{-1} \\ &= \frac{1}{3}I + w_i^* P_i \left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \alpha_i I - \alpha_i I \right) P_i^{-1} \\ &= \frac{1}{3}I + w_i^* (X_i - \alpha_i I). \end{aligned}$$

Because $\alpha_i \rightarrow 0$ and $X_i \rightarrow J$ and $Y_i^* = \frac{1}{3}I + w_i^* (X_i - \alpha_i I) \rightarrow Y^*$ we see that w_i must converge, say to w , with $\operatorname{Re} w \geq 0$ as $\operatorname{Re} w_i \geq 0$. We conclude that

$$Y = \frac{1}{3}I + wJ = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ w & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \in \hat{\partial}\alpha(J).$$

This means Y must be a regular subgradient:

$$S_2(J) \subset \hat{\partial}\alpha(J).$$

Case 3

In Case 3 the X_i satisfy

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 1 \\ 0 & 0 & \alpha_i \end{pmatrix} P_i^{-1}$$

which can be rewritten as

$$X_i = P_i \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \alpha_i \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} P_i^{-1}$$

so this case reduces to Case 2:

$$S_3(J) \subset \hat{\partial}\alpha(J).$$

Case 4

In this case we require

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 1 \\ 0 & 0 & \alpha_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ w_i & \frac{1}{3} & 0 \\ y_i & w_i & \frac{1}{3} \end{pmatrix} P_i^*.$$

Again the structure of Y_i comes from Theorem 2, so we have $\operatorname{Re} w_i \geq 0$ but $y_i \in \mathbb{C}$ is unrestricted. We know that Y satisfies (9). Since $\det(Y_i) = \frac{1}{27}$ is constant it has to be stable under the limit and that means that

$$\det(Y) = a^2(1 - 2a) = \frac{1}{27}.$$

The only solution for $a \in [0, \frac{1}{2}]$ is then $a = \frac{1}{3}$, so

$$S_4(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ b & \frac{1}{3} & c \\ d & 0 & \frac{1}{3} \end{pmatrix} \right\}.$$

Case 5

We consider two “sub-cases”.

$\operatorname{Re} \alpha_i > \operatorname{Re} \beta_i$

Here we require

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ w_i & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^*,$$

with $\operatorname{Re} w_i \geq 0$, so we have

$$\begin{aligned} X_i Y_i^* &= P_i \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} \begin{pmatrix} \frac{1}{2} w_i^* & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \\ &= P_i \begin{pmatrix} \frac{\alpha_i}{2} & \alpha_i w_i^* + \frac{1}{2} & 0 \\ 0 & \frac{\alpha_i}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} = \alpha_i Y_i^* + H_i \\ &\text{with } H_i = P_i \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1}. \end{aligned}$$

We know that $\alpha_i \rightarrow 0$ and $X_i Y_i^* \rightarrow J Y^* = aJ$ since Y satisfies (9). Therefore

$$H_i \rightarrow aJ.$$

We will show that this means that $a = \frac{1}{2}$. Consider

$$Y_i^* = \frac{1}{2}I + K_i \text{ with } K_i = P_i \begin{pmatrix} 0 & w_i^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} P_i^{-1}.$$

The sequence Y_i converges and therefore $K_i = Y_i^* - \frac{1}{2}I$ converges and so does

$$K_i^2 = P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} P_i^{-1}.$$

We have

$$X_i = \alpha_i I + 2H_i + 4(\beta_i - \alpha_i)K_i^2 \rightarrow 0 + 2aJ + 0$$

and we can conclude, since $X_i \rightarrow J$, that $J = 2aJ$, and therefore $a = \frac{1}{2}$.

Using $K_i = 2w_i^* H_i - 2K_i^2$ we see that w_i must converge, say to w , with $\operatorname{Re} w \geq 0$. We get then that

$$K_i = Y_i^* - \frac{1}{2}I \rightarrow Y^* - \frac{1}{2}I = \begin{pmatrix} 0 & b^* & d^* \\ 0 & 0 & 0 \\ 0 & c^* & -\frac{1}{2} \end{pmatrix} =: K,$$

So,

$$K_i^2 \rightarrow K^2 = \begin{pmatrix} 0 & c^* d^* & -\frac{d^*}{2} \\ 0 & 0 & 0 \\ 0 & -\frac{c^*}{2} & \frac{1}{4} \end{pmatrix}.$$

Hence,

$$K_i = 2w_i^* H_i - 2K_i^2 \rightarrow 2w^* aJ - 2K^2 = \begin{pmatrix} 0 & w^* - 2c^* d^* & d^* \\ 0 & 0 & 0 \\ 0 & c^* & -\frac{1}{2} \end{pmatrix}.$$

It follows from the equations for K that $b^* = w^* - 2c^* d^*$. Thus,

$$S_5^1(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ b & \frac{1}{2} & c \\ d & 0 & 0 \end{pmatrix} \text{ with } \operatorname{Re}(b + 2cd) \geq 0 \right\}.$$

Here the superscript denotes the sub-case of Case 5.

Re $\alpha_i = \text{Re } \beta_i$

The sequences

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

in this case are given by

$$X_i = P_i \begin{pmatrix} \alpha_i & 1 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} p_i & 0 & 0 \\ w_i & p_i & 0 \\ 0 & 0 & 1 - 2p_i \end{pmatrix} P_i^*$$

where $\text{Re } w_i \geq 0$ and $p_i \in [0, \frac{1}{2}]$, so, w.l.o.g. we can assume that p_i converges. There are two “sub-sub-cases”.

1. Suppose $p_i \rightarrow \frac{1}{3}$. Then

$$\begin{aligned} Y_i^* - p_i I &= P_i \begin{pmatrix} 0 & w_i^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 - 3p_i \end{pmatrix} P_i^{-1} \\ &\rightarrow \begin{pmatrix} a - \frac{1}{3} & b^* & d^* \\ 0 & a - \frac{1}{3} & 0 \\ 0 & c^* & \frac{2}{3} - 2a \end{pmatrix}. \end{aligned}$$

Since the determinant is a continuous function we conclude that $a = \frac{1}{3}$, so

$$S_5^{2,1}(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ b & \frac{1}{3} & c \\ d & 0 & \frac{1}{3} \end{pmatrix} \right\}.$$

Here the notation indicates that this is the first sub-sub-case of the second sub-case of Case 5.

2. Now assume $p_i \rightarrow p$ where $p \neq \frac{1}{3}$. Here we will consider

$$\begin{aligned} L_i &= Y_i^* - p_i I = P_i \begin{pmatrix} 0 & w_i^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 - 3p_i \end{pmatrix} P_i^{-1} \\ &\rightarrow Y^* - pI = \begin{pmatrix} a - p & b^* & d^* \\ 0 & a - p & 0 \\ 0 & c^* & 1 - 2a - p \end{pmatrix} \end{aligned}$$

and therefore

$$L_i^2 = P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (1 - 3p_i)^2 \end{pmatrix} P_i^{-1} \text{ converges.}$$

We deduce that $M_i = \frac{1}{(1-3p_i)^2} L_i^2$ converges. Let's call the limit M :

$$M = \frac{1}{(1-3p)^2} (Y^* - pI)^2.$$

Now define

$$N_i = P_i \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} = X_i - \alpha_i I - (\beta_i - \alpha_i) M_i \rightarrow J.$$

This implies that w_i converges to a limit w with $\operatorname{Re} w \geq 0$ since

$$Y_i^* = p_i I + (1 - 3p_i) M_i + w_i^* N_i.$$

Now consider

$$\begin{aligned} X_i Y_i^* &= P_i \begin{pmatrix} \alpha_i p_i & \alpha_i w_i^* + p_i & 0 \\ 0 & \alpha_i p_i & 0 \\ 0 & 0 & \beta_i(1 - 2p_i) \end{pmatrix} P_i^{-1} \\ &= (\alpha_i w_i^* + p_i) N_i + \alpha_i p_i I + (\beta_i(1 - 2p_i) - \alpha_i p_i) M_i \\ &\rightarrow (0 + p) J + 0 + 0 = pJ. \end{aligned}$$

Since $X_i Y_i^* \rightarrow J Y^* = aJ$ we find that $p = a$. So,

$$Y_i^* \rightarrow aI + (1 - 3a)M + w^* J = \begin{pmatrix} a & w^* + \frac{c^* d^*}{1 - 3a} & d^* \\ 0 & a & 0 \\ 0 & c^* & 1 - 2a \end{pmatrix}.$$

This means $b^* = w^* + \frac{c^* d^*}{1 - 3a}$, so we have

$$\operatorname{Re}(b - \frac{cd}{1 - 3a}) \geq 0$$

and hence

$$S_5^{2,2}(J) \subset \left\{ Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & 1 - 2a \end{pmatrix} \text{ with } a \neq \frac{1}{3} \text{ and } \operatorname{Re}(b - \frac{cd}{1 - 3a}) \geq 0 \right\}.$$

So, we write $S_5^2(J) = S_5^{2,1}(J) \cup S_5^{2,2}(J)$ and $S_5(J) = S_5^1(J) \cup S_5^2(J)$.

Case 6

We consider two sub-cases.

Re $\alpha_i > \text{Re } \beta_i$

Here the sequences

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

are given by

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^*.$$

We easily get a contradiction which rules out this case. Assume that

$$P_i = \begin{pmatrix} p_{11}^i & p_{12}^i & p_{13}^i \\ p_{21}^i & p_{22}^i & p_{23}^i \\ p_{31}^i & p_{32}^i & p_{33}^i \end{pmatrix}. \quad (12)$$

Then

$$(X_i)_{12} = \frac{p_{13}^i(\alpha_i - \beta_i)(p_{11}^i p_{32}^i - p_{12}^i p_{31}^i)}{\det(P_i)} \rightarrow 1$$

and

$$(Y_i^*)_{12} = \frac{1}{2} \frac{p_{13}^i(p_{11}^i p_{32}^i - p_{12}^i p_{31}^i)}{\det(P_i)} = \frac{1}{2} \frac{(X_i)_{12}}{\alpha_i - \beta_i}$$

which doesn't converge. So,

$$S_6^1(J) = \emptyset.$$

Re $\alpha_i = \text{Re } \beta_i$

Now we have

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \alpha_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} q_i & 0 & 0 \\ 0 & q_i & 0 \\ 0 & 0 & 1 - 2q_i \end{pmatrix} P_i^*.$$

Assume again that (12) holds, so

$$(X_i)_{12} = \frac{p_{13}^i(\alpha_i - \beta_i)(p_{11}^i p_{32}^i - p_{12}^i p_{31}^i)}{\det(P_i)} \rightarrow 1$$

and

$$(Y_i^*)_{12} = \frac{1}{2} \frac{p_{13}^i(3q_i - 1)(p_{11}^i p_{32}^i - p_{12}^i p_{31}^i)}{\det(P)} = \frac{1}{2} \frac{3q_i - 1}{\alpha_i - \beta_i} (X_i)_{21}$$

which can only converge if $q_i \rightarrow \frac{1}{3}$. But this means that $a = \frac{1}{3}$ by a determinant argument which gives:

$$S_6(J) = S_6^2(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ b & \frac{1}{3} & c \\ d & 0 & \frac{1}{3} \end{pmatrix} \right\}.$$

Case 7

We consider two sub-cases.

Re $\alpha_i > \text{Re } \beta_i$

Let

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 1 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^*$$

So

$$X_i Y_i^* = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} = \alpha_i Y_i^* \rightarrow 0.$$

Since $X_i Y_i^* \rightarrow J Y^* = a J$ we deduce that $a = 0$. Considering then that, since all the Y_i are rank one, the limit also has rank one, we have

$$Y = \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 1 \end{pmatrix} \text{ with } b = cd.$$

So,

$$S_7^1(J) \subset \left\{ Y = \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 1 \end{pmatrix} \text{ with } b = cd \right\}.$$

Re $\alpha_i = \text{Re } \beta_i$

This case is analogous to the second sub-case of Case 5 since once the real parts are the same the ordering becomes arbitrary. So, $S_7^2(J) = S_5^2(J)$ and $S_7(J) = S_7^1(J) \cup S_7^2(J)$.

Case 8

We consider two sub-cases.

Re $\alpha_i > \text{Re } \beta_i$

The discussion of this case is analogous to the first part of Case 7:

$$S_8^1(J) \subset \left\{ Y = \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 1 \end{pmatrix} \text{ with } b = cd \right\}.$$

Re $\alpha_i = \text{Re } \beta_i$

We have

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 0 \\ 0 & 0 & \beta_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} r_i & 0 & 0 \\ 0 & \frac{1}{2} - \frac{r_i}{2} & 0 \\ 0 & 0 & \frac{1}{2} - \frac{r_i}{2} \end{pmatrix} P_i^*$$

with $r_i \in [0, \frac{1}{2}]$. We have that $X_i Y_i^* \rightarrow aJ$ and

$$X_i Y_i^* = \beta_i Y_i^* + r_i P_i \begin{pmatrix} \alpha_i - \beta_i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1}$$

and

$$X_i Y_i^* = \alpha_i Y_i^* + \left(\frac{1}{2} - \frac{r_i}{2}\right) P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & \beta_i - \alpha_i & 0 \\ 0 & 0 & \beta_i - \alpha_i \end{pmatrix} P_i^{-1}$$

so since

$$P_i \begin{pmatrix} \alpha_i - \beta_i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} = X_i - \beta_i I \rightarrow J$$

and

$$P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & \beta_i - \alpha_i & 0 \\ 0 & 0 & \beta_i - \alpha_i \end{pmatrix} P_i^{-1} = X_i - \alpha_i I \rightarrow J$$

we obtain $r_i \rightarrow a$ and $\frac{1}{2} - \frac{r_i}{2} \rightarrow a$. This means that $a = \frac{1}{2} - \frac{a}{2}$ giving $a = \frac{1}{3}$. So,

$$S_8^2(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ b & \frac{1}{3} & c \\ d & 0 & \frac{1}{3} \end{pmatrix} \right\}.$$

Case 9

We consider three sub-cases.

Re $\alpha_i > \text{Re } \beta_i \geq \text{Re } \gamma_i$

We follow the same argument as in the first part of Case 7. Thus

$$S_9^1(J) \subset \left\{ Y = \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 1 \end{pmatrix} \text{ with } b = cd \right\}.$$

$\operatorname{Re} \alpha_i = \operatorname{Re} \beta_i > \operatorname{Re} \gamma_i$

We have $X_i \rightarrow J$ with

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 0 \\ 0 & 0 & \gamma_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} r_i & 0 & 0 \\ 0 & 1 - r_i & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^*$$

Since $\det(Y_i)=0$ we conclude that $\det(Y) = 0$ and therefore $a = \frac{1}{2}$ or $a = 0$. Since $r_i \in [0, \frac{1}{2}]$, w.l.o.g. we assume that $r_i \rightarrow r$.

1. We first consider the case $a = \frac{1}{2}$. We know that $\det(Y_i - (1 - r_i)I) = 0$. By taking the limit we get that $\det(Y - (1 - r)I) = 0$ from which we deduce that $r = \frac{1}{2}$. Consider $(Y_i^* - r_i I)Y_i^*$ and $(X_i - \alpha_i I)Y_i^*$:

$$\begin{aligned} (Y_i^* - r_i I)Y_i^* &= P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (1 - 2r_i)(1 - r_i) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \\ &\rightarrow (Y^* - \frac{1}{2}I)Y^* = \begin{pmatrix} 0 & \frac{b^*}{2} + c^*d^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ (X_i - \alpha_i I)Y_i^* &= P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\beta_i - \alpha_i)(1 - r_i) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \\ &\rightarrow JY^* = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Defining Q_i such that

$$\begin{aligned} (Y_i^* - r_i I)Y_i^* &= (1 - 2r_i)Q_i \\ (X_i - \alpha_i I)Y_i^* &= (\beta_i - \alpha_i)Q_i \end{aligned}$$

we get by looking at the quotient of the (1,2) entries that

$$\frac{1 - 2r_i}{\beta_i - \alpha_i} \rightarrow \frac{b^*/2 + c^*d^*}{1/2} = b^* + 2c^*d^*.$$

Since $\operatorname{Re} \alpha_i = \operatorname{Re} \beta_i$ we get that $\operatorname{Re}(\frac{1-2r_i}{\beta_i - \alpha_i}) = 0$ and conclude that $\operatorname{Re}(b + 2cd) = 0$. So,

$$S_9^{2,1}(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ b & \frac{1}{2} & c \\ d & 0 & 0 \end{pmatrix} \text{ with } \operatorname{Re}(b + 2cd) = 0 \right\}.$$

2. Now we consider the case $a = 0$. Since $\det(Y_i - r_i I) = 0$ and $r_i \in [0, \frac{1}{2}]$ we get $r = 0$. Consider $(Y_i^* - (1 - r_i)I)Y_i^*$ and $(Y_i^* - (1 - r_i)I)(X_i - \gamma_i I)$:

$$\begin{aligned} (Y_i^* - (1 - r_i)I)Y_i^* &= P_i \begin{pmatrix} (2r_i - 1)r_i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \\ &\rightarrow (Y^* - I)Y^* = \begin{pmatrix} 0 & c^* d^* - b^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ (Y_i^* - (1 - r_i)I)(X_i - \gamma_i I) &= P_i \begin{pmatrix} (2r_i - 1)(\alpha_i - \gamma_i) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \\ &\rightarrow (Y^* - I)J = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Following a similar argument as in the previous sub-sub-case, we get that $\frac{r_i}{\alpha_i - \gamma_i} \rightarrow b^* - c^* d^*$ and therefore $\operatorname{Re}(b - cd) \geq 0$. So,

$$S_9^{2,2}(J) \subset \left\{ Y = \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 1 \end{pmatrix} \text{ with } \operatorname{Re}(b - cd) \geq 0 \right\}$$

$$\underline{\operatorname{Re} \alpha_i = \operatorname{Re} \beta_i = \operatorname{Re} \gamma_i}$$

Let

$$X_i \rightarrow J \text{ and } Y_i \rightarrow Y$$

where

$$X_i = P_i \begin{pmatrix} \alpha_i & 0 & 0 \\ 0 & \beta_i & 0 \\ 0 & 0 & \gamma_i \end{pmatrix} P_i^{-1} \text{ and } Y_i = P_i^{-*} \begin{pmatrix} r_i & 0 & 0 \\ 0 & q_i & 0 \\ 0 & 0 & 1 - r_i - q_i \end{pmatrix} P_i^*.$$

We can assume that w.l.o.g $r_i \rightarrow r$ and $q_i \rightarrow q$. The Cayley-Hamilton theorem says that

$$0 = (Y_i - r_i I)(Y_i - q_i I)(Y_i - (1 - q_i - r_i)I)$$

so

$$0 = (Y - rI)(Y - qI)(Y - (1 - q - r)I).$$

Looking at the diagonal entries in this matrix product we see that one of the following 4 sub-sub-cases must hold:

1. $a \neq \frac{1}{3}$ and $q = a$ and $r = a$
2. $a \neq \frac{1}{3}$ and $q = a$ and $r = 1 - 2a$
3. $a \neq \frac{1}{3}$ and $q = 1 - 2a$ and $r = a$
4. $a = q = r = \frac{1}{3}$

We consider these in turn.

1. Consider $(Y_i^* - r_i I)(Y_i^* - (1 - q_i - r_i)I)$. This gives

$$P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (q_i - r_i)(2q_i + r_i - 1) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \rightarrow \begin{pmatrix} 0 & b^*(3a - 1) + c^*d^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Looking at $(X_i - \alpha_i I)(Y_i^* - (1 - r_i - q_i)I)$ we get that

$$P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\beta_i - \alpha_i)(2q_i + r_i - 1) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \rightarrow \begin{pmatrix} 0 & 3a - 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Comparing the quotient as above we can deduce that

$$\frac{q_i - r_i}{\beta_i - \alpha_i} \rightarrow b^* + \frac{c^*d^*}{3a - 1}.$$

Since q_i and r_i are real and $\beta_i - \alpha_i$ is imaginary we get that

$$\operatorname{Re}(b + \frac{cd}{3a - 1}) = 0,$$

so

$$S_9^{3,1}(J) \subset \left\{ Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & 1 - 2a \end{pmatrix} \text{ with } \operatorname{Re}(b - \frac{cd}{1 - 3a}) = 0 \right\}.$$

2. Consider $(Y_i^* - r_i I)(Y_i^* - (1 - q_i - r_i)I)$. This gives

$$P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (q_i - r_i)(2q_i + r_i - 1) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \rightarrow \begin{pmatrix} 0 & b^*(3a - 1) + c^*d^* & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Looking at $(X_i - \gamma_i I)(Y_i^* - r_i I)$ we get that

$$P_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\beta_i - \gamma_i)(q_i - r_i) & 0 \\ 0 & 0 & 0 \end{pmatrix} P_i^{-1} \rightarrow \begin{pmatrix} 0 & 3a - 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Comparing the quotient again we can deduce that

$$\frac{2q_i + r_i - 1}{\beta_i - \gamma_i} \rightarrow b^* + \frac{c^*d^*}{3a - 1}.$$

Since q_i and r_i are real and $\beta_i - \gamma_i$ is imaginary we get

$$\operatorname{Re}(b - \frac{cd}{1 - 3a}) = 0,$$

so

$$S_9^{3,2}(J) \subset \left\{ Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & 1 - 2a \end{pmatrix} \text{ with } \operatorname{Re}(b - \frac{cd}{1 - 3a}) = 0 \right\}.$$

3. Since q_i and r_i are interchangeable, this is the same as the previous sub-case.
4. Because $a = \frac{1}{3}$ we get

$$S_9^{3,4}(J) \subset \left\{ Y = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ b & \frac{1}{3} & c \\ d & 0 & \frac{1}{3} \end{pmatrix} \right\}.$$

Finally,

$$S_9(J) = S_9^1(J) \cup S_9^{2,1}(J) \cup S_9^{2,2}(J) \cup S_9^{3,1}(J) \cup S_9^{3,2}(J) \cup S_9^{3,4}(J).$$

Thus, after considering all possible limits we conclude that all $S_j(J) \subset \gamma(J)$ which proves that $\partial\alpha(J) \subset \gamma(J)$. In fact, we showed that $S_j(J)$ is strictly contained in $\gamma(J)$ except when $j = 5$ or $j = 7$. Note that these are the two cases where X_i has precisely two nonderogatory eigenvalues.

5 Proof of the inclusion $\partial\alpha(J) \supset \gamma(J)$

We want to prove that any $Y \in \gamma(J)$ is a subgradient of α at J . We will distinguish the cases $a = \frac{1}{3}$ and $a \neq \frac{1}{3}$.

The case $a = \frac{1}{3}$.

First suppose that c and d are nonzero. The sequences

$$\begin{aligned} P_t &= \begin{pmatrix} -c^*d^{*2} & -\frac{b^*d^*}{t} & \frac{c^*d^{*2}}{t^3} \\ 0 & -c^*d^{*2} & 0 \\ 0 & -\frac{c^*d^*}{t} & \frac{1}{t^3} \end{pmatrix} \\ P_t^{-1} &= \begin{pmatrix} -\frac{1}{c^*d^{*2}} & \frac{b^*-c^{*2}d^{*2}}{c^{*2}d^{*3}t} & 1 \\ 0 & -\frac{1}{c^*d^{*2}} & 0 \\ 0 & -\frac{t^2}{d^*} & t^3 \end{pmatrix} \\ J_t &= \begin{pmatrix} t & 1 & 0 \\ 0 & t & 1 \\ 0 & 0 & t \end{pmatrix} \\ W_t^* &= \begin{pmatrix} \frac{1}{3} & \frac{1}{t} & -\frac{1}{c^*d^{*3}t^3} \\ 0 & \frac{1}{3} & \frac{1}{t} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}. \end{aligned}$$

are such that $X_t = P_t J_t P_t^{-1}$, $Y_t = P_t^{-*} W_t P_t^* \in \hat{\partial}\alpha(X_t)$,

$$X_t \rightarrow J \text{ and } Y_t \rightarrow Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & a \end{pmatrix} \text{ with } a = \frac{1}{3}.$$

Here t is real and positive and decreases to zero. If either c or d is zero, we simply replace it where it occurs in these sequences by t . Thus, as desired, we created a sequence for $a = \frac{1}{3}$ and arbitrary b, c, d for which $Y_t \in \hat{\partial}\alpha(X_t)$ converges to the specified $Y \in \gamma(J)$.

The case $a \neq \frac{1}{3}$.

We create a sequence for arbitrary $a \in [0, \frac{1}{2})$ and $a \neq \frac{1}{3}$ (we will discuss the case $a = \frac{1}{2}$ later) and c, d arbitrary and $\operatorname{Re} w \geq 0$:

$$\begin{aligned} P_t^{-1} &= \begin{pmatrix} \frac{1}{t} & \frac{1}{t} & \frac{c^*}{(3a-1)t} \\ t & \frac{1}{t} & t \\ t & \frac{d^*}{(1-3a)t} & \frac{1}{t} \end{pmatrix}, \\ J_t &= \begin{pmatrix} t & 1 & 0 \\ 0 & t & 0 \\ 0 & 0 & t + it \end{pmatrix}, \\ W_t^* &= \begin{pmatrix} a + t & w^* + t & 0 \\ 0 & a + t & 0 \\ 0 & 0 & 1 - 2a - 2t \end{pmatrix} \end{aligned} \quad (13)$$

are such that $X_t = P_t^{-1} J_t P_t$, $Y_t = P_t^* W_t P_t^{-*} \in \hat{\partial}\alpha(X_t)$,

$$X_t \rightarrow J \text{ and } Y_t \rightarrow Y = \begin{pmatrix} a & 0 & 0 \\ b & a & c \\ d & 0 & a \end{pmatrix}$$

where $b = w + \frac{cd}{1-3a}$, with t positive real, converging to 0. We omit the formula for P_t since it is complicated. In the case where $a = \frac{1}{2}$ we can take a similar sequence with

$$W_t^* = \begin{pmatrix} \frac{1}{2} - t & w^* + t & 0 \\ 0 & \frac{1}{2} - t & 0 \\ 0 & 0 & 2t \end{pmatrix}.$$

We then get that $X_t \rightarrow J$ and $Y_t \rightarrow Y$ as above but with $a = \frac{1}{2}$. It is necessary to consider this separately since in equation (13) if $a = \frac{1}{2}$ the bottom right entry would be negative and therefore not a regular subgradient. Since $b = w + \frac{cd}{1-3a}$ and $\operatorname{Re} w \geq 0$ we created a sequence for $a \neq \frac{1}{3}$ and arbitrary b, c, d satisfying $\operatorname{Re}(b - \frac{cd}{1-3a})$ for which $Y_t \in \hat{\partial}\alpha(X_t)$ converges to the specified $Y \in \gamma(J)$.

This proves the inclusion $\partial\alpha(J) \supset \gamma(J)$ and completes the proof of Theorem 5.

6 The horizon subgradients

We conclude our subdifferential analysis for the spectral abscissa at J by giving necessary and sufficient conditions for the horizon subgradients. Note that the expression for $\partial^\infty \alpha(J)$ is simpler than the expression for $\partial \alpha(J)$ because there is no inequality constraint on the matrix entries.

Theorem 6

$$\partial^\infty \alpha(J) = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 0 \end{pmatrix} \right\}.$$

Proof. The necessary conditions for the horizon subgradient are proved by applying Theorem 1 so all we need to do is establish sufficient conditions. Consider the following sequences of matrices:

$$\begin{aligned} P_t &= \begin{pmatrix} -c^* d^{*2} & -\frac{b^* d^*}{t} & \frac{c^* d^{*2}}{t^3} \\ 0 & -c^* d^{*2} & 0 \\ 0 & -\frac{c^* d^*}{t} & \frac{1}{t^3} \end{pmatrix} \\ P_t^{-1} &= \begin{pmatrix} -\frac{1}{c^* d^{*2}} & \frac{b^* - c^{*2} d^{*2}}{c^{*2} d^{*3} t} & 1 \\ 0 & -\frac{1}{c^* d^{*2}} & 0 \\ 0 & -\frac{t^2}{d^*} & t^3 \end{pmatrix} \\ J_t &= \begin{pmatrix} t & 1 & 0 \\ 0 & t & 1 \\ 0 & 0 & t \end{pmatrix} \\ W_t^* &= \begin{pmatrix} \frac{1}{3} & \frac{1}{t^2} & -\frac{1}{c^* d^{*4} t^4} \\ 0 & \frac{1}{3} & \frac{1}{t^2} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}. \end{aligned}$$

They have the property that $X_t = P_t J_t P_t^{-1} \rightarrow J$ and with $Y_t = P_t^* W_t P_t^{-*} \in \partial \hat{\alpha}(X_t)$, we have

$$tY_t \rightarrow \begin{pmatrix} 0 & 0 & 0 \\ b & 0 & c \\ d & 0 & 0 \end{pmatrix}$$

for any choice of $b, c \neq 0, d \neq 0$ and t positive real, converging to 0. If either c or d is zero change it to t in the sequence above. \blacksquare

7 Concluding remarks

The subgradient analysis of Sections 4 and 5, establishing necessary and sufficient conditions for Y to be a subgradient of the spectral abscissa at the 3×3 derogatory, defective matrix J , was quite complicated, suggesting that a general analysis for derogatory, defective matrices may be difficult. However, we hope that the detailed analysis presented for this case may show the way forward for obtaining a more general result.

References

- [Arn71] V.I. Arnold. On matrices depending on parameters. *Russian Mathematical Surveys*, 26:29–43, 1971.
- [BHLO06] J.V. Burke, D. Henrion, A.S. Lewis, and M.L. Overton. Stabilization via non-smooth, nonconvex optimization. *IEEE Transactions on Automatics Control*, 51:1760–1769, 2006.
- [BLO01] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimizing matrix stability. *Proceedings of the American Mathematical Society*, 129:1635–1642, 2001.
- [BO01] J.V. Burke and M.L. Overton. Variational analysis of non-Lipschitz spectral functions. *Mathematical Programming*, 90:317–352, 2001.
- [GO07] K. K. Gade and M. L. Overton. Optimizing the asymptotic convergence rate of the Diaconis-Holmes-Neal sampler. *Adv. in Appl. Math.*, 38(3):382–403, 2007.
- [Gru11] S. Grundel. *Eigenvalue Optimization in C^2 Subdivision and Boundary Subdivision*. PhD thesis, New York University, 2011. <http://cs.nyu.edu/overton/phdtheses/sara.pdf>.
- [HO06] D. Henrion and M. L. Overton. Maximizing the closed loop asymptotic decay rate for the two-mass-spring control problem. Technical Report 06342, LAAS-CNRS, March 2006. <http://homepages.laas.fr/henrion/Papers/massspring.pdf>.
- [Mor76] B.S. Mordukhovich. Maximum principle in the problem of time optimal response with nonsmooth constraints. *Journal of Applied Mathematics and Mechanics*, 40:960–969, 1976.
- [OW88] M.L. Overton and R.S. Womersley. On minimizing the spectral radius of a non-symmetric matrix function – optimality conditions and duality theory. *SIAM Journal on Matrix Analysis and Applications*, 9:473–498, 1988.
- [RW98] R.T. Rockafellar and R.J.B. Wets. *Variational Analysis*. Springer, New York, 1998.