μ-VALUES AND SPECTRAL VALUE SETS FOR LINEAR PERTURBATION CLASSES DEFINED BY A SCALAR PRODUCT

MICHAEL KAROW *

Abstract. We study the variation of the spectrum of matrices under perturbations which are self- or skew-adjoint with respect to a scalar product. Computable formulae are given for the associated μ-values. The results can be used to calculate spectral value sets for the perturbation classes under consideration. We discuss the special case of complex Hamiltonian perturbations of a Hamiltonian matrix in detail.

Key words. linear systems, eigenvalues, perturbations, spectral value sets, μ-values

AMS subject classifications. 15A18, 15A57, 15A63, 93C05, 93C73

1. Introduction. μ-values are well established tools in stability analysis of uncertain systems and in eigenvalue perturbation theory [10, 13, 22, 27]. They can be used to characterize several important quantities including stability radii, structured eigenvalue condition numbers [14] and the structured distance to uncontrollability [16]. The relationship of spectral value sets with μ-values will be displayed below. There is a vast literature on the problem of calculating μ-values with respect to various perturbation classes [1, 3, 5, 15, 23, 24]. In this paper we give computable formulae for μ if the underlying perturbation class is a set of self-adjoint or skew-adjoint matrices with respect to a scalar product. The scalar product is assumed to be defined by a unitary matrix, see Section 6. It will be shown that in this case the associated μ-values can be obtained by solving a simple one parameter optimization problem.

We use the following notation. The symbols $\mathbb{N}, \mathbb{R}, \mathbb{C}$ represent the sets of positive integers, real numbers and complex numbers respectively. By $\mathbb{C}^{n\times m}$ we denote the set of n by m matrices with entries in $\mathbb{C}$. Furthermore, $\mathbb{C}^n = \mathbb{C}^{n\times 1}$ is the set of column vectors of length n. The conjugate, the transpose and the conjugate transpose of $A \in \mathbb{C}^{n\times m}$ will be written $\overline{A}$, $A^\top$ and A^*. If A is square then $\sigma(A)$ and $\rho(A) = \mathbb{C} \setminus \sigma(A)$ denote its spectrum and its resolvent set. The $n \times n$ identity matrix will be written I_n. By a perturbation class $\mathbf{\Delta}$ we mean a nonempty closed subset of $\mathbb{C}^{l\times q}$ which is star shaped with respect to $0 \in \mathbb{C}^{l\times q}$, i.e. if $\Delta \in \mathbf{\Delta}$ then $t\Delta \in \mathbf{\Delta}$ for $0 \le t \le 1$. We now give the definition of μ-values.

DEFINITION 1.1. *Let $\mathbf{\Delta} \subseteq \mathbb{C}^{l\times q}$ be a perturbation class and let $\|\cdot\|$ be a norm on $\mathbb{C}^{l\times q}$.*

- *The μ-value of $M \in \mathbb{C}^{q\times l}$ with respect to $\mathbf{\Delta}$ and $\|\cdot\|$ is*

$$\mu_{\mathbf{\Delta}}(M) := \left(\inf\{\, \|\Delta\|\,;\, \Delta \in \mathbf{\Delta},\ 1 \in \sigma(\Delta M)\,\}\,\right)^{-1}. \tag{1.1}$$

  Thus $\mu_{\mathbf{\Delta}}(M)$ is the inverse of the smallest norm of a $\Delta \in \mathbf{\Delta}$ such that 1 is an eigenvalue of the matrix product ΔM. If there is no such $\Delta \in \mathbf{\Delta}$ then $\mu_{\mathbf{\Delta}}(M) = 0$.
- *If $l = q$ then the μ-value of M of second kind is defined as*

$$\widetilde{\mu}_{\mathbf{\Delta}}(M) := \inf\{\, \|\Delta\|\,;\ \Delta \in \mathbf{\Delta},\ \det(M - \Delta) = 0\,\}. \tag{1.2}$$

  Thus $\widetilde{\mu}_{\mathbf{\Delta}}(M)$ is the structured distance of M to the set of singular matrices. We have $\widetilde{\mu}_{\mathbf{\Delta}}(M) = 0$ iff M is singular, and $\widetilde{\mu}_{\mathbf{\Delta}}(M) = \infty$ iff there is no $\Delta \in \mathbf{\Delta}$ such that $\det(M - \Delta) = 0$.

It is easily seen that $\widetilde{\mu}_{\mathbf{\Delta}}(M) = \mu_{\mathbf{\Delta}}(M^{-1})^{-1}$ if M is nonsingular. Furthermore, if the underlying norm is the spectral norm then

$$\mu_{\mathbb{C}^{l\times q}}(M) = \sigma_{\max}(M), \qquad \text{and} \qquad \widetilde{\mu}_{\mathbb{C}^{n\times n}}(M) = \sigma_{\min}(M), \tag{1.3}$$

where $\sigma_{\max}(\cdot)$ and $\sigma_{\min}(\cdot)$ denote the maximum and the minimum singular value respectively.

We now briefly discuss the relationship of μ-values with the perturbation analysis of eigenvalues. Consider matrix perturbations of the form

$$A \rightsquigarrow A_\Delta = A + B\Delta C, \qquad \Delta \in \mathbf{\Delta}, \qquad \|\Delta\| < \delta, \tag{1.4}$$

where $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times l}$, $C \in \mathbb{C}^{q\times n}$ are fixed matrices. The set of all eigenvalues of all matrices A_Δ given by (1.4) is called a spectral value set (stuctured pseudospectrum). It is denoted by

$$\begin{aligned}\sigma_{\mathbf{\Delta}}(A, B, C; \delta) &:= \bigcup_{\Delta\in\mathbf{\Delta},\ \|\Delta\|<\delta} \sigma(A + B\Delta C) \\ &= \{s \in \mathbb{C};\ \exists\, \Delta \in \mathbf{\Delta} : \|\Delta\| < \delta, \text{ and } \det(sI_n - (A + B\Delta C)) = 0\}.\end{aligned} \tag{1.5}$$

*Mathematics Institute, Berlin University of Technology, D-10623 Berlin, Germany, (karow@math.TU-Berlin.de).

Let $G(s) := C(sI_n - A)^{-1}B,\ s \in \rho(A)$, be the transfer function of the triple (A, B, C). From the well known equivalence [9, Proposition 2.3]

$$s\in\sigma(A + B\Delta C) \ \Leftrightarrow\ 1\in\sigma(\Delta G(s)) \tag{1.6}$$

it follows that

$$\mu_{\boldsymbol{\Delta}}(G(s)) = (\,\inf\{\,\|\Delta\|\,;\ \ \Delta\in\boldsymbol{\Delta},\ s\in\sigma(A + B\Delta C)\,\}\,)^{-1}, \quad s\in\rho(A). \tag{1.7}$$

This in turn yields

$$\sigma_{\boldsymbol{\Delta}}(A, B, C; \delta) = \sigma(A) \cup \{s \in \rho(A);\ \mu_{\boldsymbol{\Delta}}(G(s)) > \delta^{-1}\}, \quad \delta > 0. \tag{1.8}$$

For the case $B = C = I_n$ and $\boldsymbol{\Delta} \subseteq \mathbb{C}^{n\times n}$ we simplify notation and denote the associated spectral value sets by

$$\sigma_{\boldsymbol{\Delta}}(A; \delta) := \sigma_{\boldsymbol{\Delta}}(A, I_n, I_n; \delta) = \bigcup_{\Delta\in\boldsymbol{\Delta},\ \|\Delta\|<\delta} \sigma(A + \Delta). \tag{1.9}$$

From the definition of $\widetilde{\mu}$ it is immediate that for $A \in \mathbb{C}^{n\times n}$,

$$\widetilde{\mu}_{\boldsymbol{\Delta}}(s\,I_n - A) = \inf\{\|\Delta\| \mid\ \Delta\in\boldsymbol{\Delta},\ s\in\sigma(A + \Delta)\,\}, \quad s\in\mathbb{C}, \tag{1.10}$$
$$\sigma_{\boldsymbol{\Delta}}(A; \delta) = \{s \in \mathbb{C};\ \widetilde{\mu}_{\boldsymbol{\Delta}}(s\,I_n - A) < \delta\,\}, \quad \delta > 0. \tag{1.11}$$

The statements (1.8) and (1.11) yield that spectral value sets can be calculated by evaluating the functions $s \mapsto \mu_{\boldsymbol{\Delta}}(G(s))$ and $s \mapsto \widetilde{\mu}_{\boldsymbol{\Delta}}(s\,I_n - A)$ respectively.

The organization of this paper is as follows. In Section 2 we provide useful characterizations for μ with respect to Hermitian, symmetric and skew-symmetric perturbations. These characterizations are then used in Sections 4 and 5 to compute the associated μ-values by maximizing or minimizing a certain eigenvalue of a Hermitian pencil. Some facts on Hermitian matrices which are needed in the proofs are given in Section 3. In Section 6 we show how the results obtained so far can be extended to the perturbation classes of self- and skew-adjoint matrices with respect to a scalar product. The last section deals with a special case: μ-values and spectral value sets for Hamiltonian perturbations of Hamiltonian matrices.

2. Hermitian, symmetric, and skew-symmetric perturbations . In this section we consider μ-values with respect to the perturbation classes $\boldsymbol{\Delta} \in \{Herm(n), Sym(n), Skew(n)\}$, where

$$\begin{array}{rcl} Herm(n) & := & \{\,\Delta \in \mathbb{C}^{n\times n}\ ;\ \Delta^* = \Delta\,\}, \\ Sym(n) & := & \{\,\Delta \in \mathbb{C}^{n\times n}\ ;\ \Delta^\top = \Delta\,\}, \\ Skew(n) & := & \{\,\Delta \in \mathbb{C}^{n\times n}\ ;\ \Delta^\top = -\Delta\,\}. \end{array} \tag{2.1}$$

First, we give a characterization of μ which holds for arbitrary perturbation classes $\boldsymbol{\Delta} \subseteq \mathbb{C}^{l\times q}$. Let

$$\nu_{\boldsymbol{\Delta}}(x, y) := \inf\{\,\|\Delta\|\ ;\ \Delta \in \boldsymbol{\Delta},\ \Delta x = y\,\}, \qquad x \in \mathbb{C}^q, y \in \mathbb{C}^l.$$

Note that

$$\nu_{\boldsymbol{\Delta}}(x, y) \ge \|y\|/\|x\| \text{ for all } x \ne 0, \tag{2.2}$$

since $\Delta x = y$ implies $\|\Delta\|\|x\| \ge \|y\|$.

LEMMA 2.1. *For any $M \in \mathbb{C}^{q\times l}$,*

$$\mu_{\boldsymbol{\Delta}}(M) = (\inf\{\,\nu_{\boldsymbol{\Delta}}(Mv, v)\ ;\ \ v\in\mathbb{C}^n,\ \ \|v\| = 1\ \ \})^{-1}.$$

If $M \in \mathbb{C}^{n\times n}$ and $\boldsymbol{\Delta} \subseteq \mathbb{C}^{n\times n}$ then

$$\widetilde{\mu}_{\boldsymbol{\Delta}}(M) = \inf\{\,\nu_{\boldsymbol{\Delta}}(v, Mv)\ ;\ \ v\in\mathbb{C}^n,\ \ \|v\| = 1\ \ \}.$$

Proof. This follows from the equivalences

$$\begin{aligned} 1 \in \sigma(\Delta M) &\Leftrightarrow \Delta(Mv) = v \text{ for some } v \text{ with } \|v\| = 1, \\ \det(M - \Delta) = 0 &\Leftrightarrow \Delta v = Mv \text{ for some } v \text{ with } \|v\| = 1. \end{aligned}$$

□

Throughout the rest of this paper the underlying norm $\|\cdot\|$ is the spectral norm. The proposition below gives the ν-values for the classes defined in (2.1).

PROPOSITION 2.1. *Let $x, y \in \mathbb{C}^n$, $x \neq 0$. Then*

(*a*) $\nu_{Herm}(x,y) = \begin{cases} \|y\|/\|x\| & \text{if } y^*x \in \mathbb{R}, \\ \infty & \text{otherwise}, \end{cases}$

(*b*) $\nu_{Skew}(x,y) = \begin{cases} \|y\|/\|x\| & \text{if } y^Tx = 0, \\ \infty & \text{otherwise}, \end{cases}$

(*c*) $\nu_{Sym}(x,y) = \|y\|/\|x\|$.

Proof. Without loss of generality we may assume that $\|x\| = 1$.

(*a*). Let $\alpha := x^*y$. Suppose there exists $\Delta \in Herm(n)$ such that $\Delta x = y$. Then $\alpha = x^*\Delta x \in \mathbb{R}$. Thus, $\nu_{Herm}(x,y) = \infty$ if $\alpha \notin \mathbb{R}$. Assume now that $\alpha \in \mathbb{R}$ and the vectors x and y are linearly independent. Let $\beta := \|y - \alpha x\|$ and $z := \beta^{-1}(y - \alpha x)$. Then $\|z\| = 1$, $x^*z = 0$, $y = \alpha x + \beta z$ and $\alpha^2 + \beta^2 = \|y\|^2$. Let

$$\Delta_0 := \alpha x x^* + \beta(z x^* + x z^*) - \alpha z z^*.$$

Then $\Delta_0 \in Herm(n)$ and $\Delta_0 x = y$. A straightforward computation yields that $\Delta_0^*\Delta_0 = \Delta_0^2 = \|y\|^2(xx^* + zz^*)$. Since the unit vectors x and z are orthogonal to each other we have $\|xx^* + zz^*\| = 1$. Hence $\|\Delta_0\| = \sqrt{\|\Delta_0^*\Delta_0\|} = \|y\|$ and therefore $\nu_{Herm}(x,y) = \|y\|$, using (2.2). Suppose now that $\alpha \in \mathbb{R}$ and x and y are linearly dependent, and set $\Delta_0 := \alpha x x^*$. Then $y = \alpha x$, $\Delta_0 \in Herm(n)$, $\Delta_0 x = y$ and $\|\Delta_0\| = |\alpha| = \|y\|$. Thus $\nu_{Herm}(x,y) = \|y\|$.

(*c*). Let $\alpha := x^Ty$. Suppose first that $\bar{x}$ (the conjugate of x) and y are linearly independent, and let $z = \beta^{-1}(\bar{y} - \bar{\alpha}x)$, where $\beta := \|\bar{y} - \bar{\alpha}x\| \neq 0$. Then we have $\|\bar{z}\| = \|z\| = 1$, $x^*z = 0 = \bar{x}^*\bar{z}$, $y = \alpha\bar{x} + \beta\bar{z}$ and $|\alpha|^2 + \beta^2 = \|y\|^2$. Set

$$\Delta_0 := \alpha \bar{x} x^* + \beta(\bar{z} x^* + \bar{x} z^*) - \bar{\alpha} \bar{z} z^*.$$

Then $\Delta_0 \in Sym(n)$ and $\Delta_0 x = y$. By a straightforward computation one obtains $\Delta_0^*\Delta_0 = \|y\|^2(xx^* + zz^*)$. Thus $\|\Delta_0\| = \|y\|$. Suppose now that $\bar{x}$ and y are linearly dependent, and set $\Delta_0 := \alpha\bar{x}x^*$. Then $y = \alpha\bar{x}$, $\Delta_0 \in Sym(n)$, $\Delta_0 x = y$ and $\|\Delta_0\| = |\alpha| = \|y\|$.

(*b*). If there exists $\Delta \in Skew(n)$ such that $\Delta x = y$ then $x^Ty = x^T\Delta x = 0$. Suppose the latter condition holds. Then the skew-symmetric matrix

$$\Delta_0 := y x^* - \bar{x} y^T$$

satisfies $\Delta_0 x = y$ and $\|\Delta_0\| = \|y\|$. □

The statement of Proposition 2.1 is covered by the results in [20]. We have given a proof here for the convenience of the reader.

By combining Proposition 2.1 with Lemma 2.1 we obtain the Theorem below.

THEOREM 2.2. *Let $M \in \mathbb{C}^{n\times n}$. Then the following holds.*

(*a*) *If the Hermitian matrix $M_h = \frac{1}{2i}(M - M^*)$ is definite then $\mu_{Herm}(M) = 0$ and $\widetilde{\mu}_{Herm}(M) = \infty$. Otherwise*

$$\begin{aligned} \mu_{Herm}(M) &= \max\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1,\ v^*M_hv = 0\ \}, \qquad (2.3) \\ \widetilde{\mu}_{Herm}(M) &= \min\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1,\ v^*M_hv = 0\ \}. \end{aligned}$$

(*b*) *Let $M_s = \frac{1}{2}(M + M^\top)$. Then for $n \geq 2$,*

$$\mu_{Skew}(M) = \max\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1,\ v^TM_sv = 0\ \} \qquad (2.4)$$

$$\widetilde{\mu}_{Skew}(M) = \min\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1,\ v^TM_sv = 0\ \}. \qquad (2.5)$$

(*c*) *We always have*

$$\begin{aligned} \mu_{Sym}(M) &= \max\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1\ \} = \sigma_{\max}(M), \\ \widetilde{\mu}_{Sym}(M) &= \min\{\ \|Mv\|\ ;\ \ v \in \mathbb{C}^n,\ \|v\| = 1\ \} = \sigma_{\min}(M). \end{aligned}$$

Proof. By (2.3) and Proposition 2.1,

$$\mu_{Herm}(M) = (\inf\{ \nu_{Herm}(Mv,v) \; ; \; v\in\mathbb{C}^n, \; \|v\| = 1 \})^{-1}, \tag{2.6}$$

$$\nu_{Herm}(Mv,v) = \begin{cases} \|Mv\|^{-1} & \text{if } 0 = \Im((Mv)^*v) = v^*M_hv \\ \infty & \textit{otherwise.} \end{cases} \quad \text{if } \|v\| = 1. \tag{2.7}$$

Hence $\mu_{Herm}(M) = \infty^{-1} = 0$ if M_h is definite. Otherwise (2.6) and (2.7) yield (2.3). The proof of the other statements is analogous. For (2.4) and (2.5) one needs the fact that $v^\top M_s v = 0$ for some nonzero $v \in \mathbb{C}^n$ if $n \geq 2$ (see Lemma 5.3). Hence $\mu_{Skew}(M) \neq \infty \neq \widetilde{\mu}_{Skew}(M)$. $\square$

Note that the μ-values for the symmetric case coincide with the μ-values for the unstructured case $\mathbf{\Delta} = \mathbb{C}^{n\times n}$ (see relation (1.3)). For the sake of computation of μ-values for the Hermitian and the skew-symmetric case we provide now a reformulation of the characterizations in Theorem 2.2. To this end we introduce the following notation. For $H, H_0, H_1 \in Herm(n)$, $S \in Sym(n)$ we define

$$\begin{array}{rcl} \overline{m}_h(H_0,H_1) & := & \sup\{ v^*H_0v \; ; \; v \in \mathbb{C}^n, \; v^*H_1v = 0, \; \|v\| = 1 \}, \\ \underline{m}_h(H_0,H_1) & := & \inf\{ v^*H_0v \; ; \; v \in \mathbb{C}^n, \; v^*H_1v = 0, \; \|v\| = 1 \}, \\ \overline{m}_{hs}(H,S) & := & \sup\{ v^*Hv \; ; \; v \in \mathbb{C}^n, \; v^TSv = 0, \; \|v\| = 1 \}, \\ \underline{m}_{hs}(H,S) & := & \inf\{ v^*Hv \; ; \; v \in \mathbb{C}^n, \; v^TSv = 0, \; \|v\| = 1 \}. \end{array} \tag{2.8}$$

We have the following corollary to Theorem 2.2.

COROLLARY 2.3. *Let $M\in\mathbb{C}^{n\times n}$, $M_h = \frac{1}{2i}(M - M^*)$, $M_s = \frac{1}{2}(M + M^\top)$. Then*

(a) if M_h is not definite,

$$\mu_{Herm}(M) = \sqrt{\overline{m}_h\,(M^*M, M_h)},$$

$$\widetilde{\mu}_{Herm}(M) = \sqrt{\underline{m}_h\,(M^*M, M_h)};$$

(b) if $n \geq 2$,

$$\mu_{Skew}(M) = \sqrt{\overline{m}_{hs}\,(M^*M, M_s)},$$

$$\widetilde{\mu}_{Skew}(M) = \sqrt{\underline{m}_{hs}\,(M^*M, M_s)}.$$

Thus, in order to calculate the μ-values for the Hermitian and the skew-symmetric case it remains to give computable formulas for the quantities defined in (2.8). This is done in the following sections.

3. Some facts on Hermitian matrices. This section contains results on Hermitian matrices which are needed later on. In the sequel $\lambda_1(H) \geq \lambda_2(H) \geq \ldots \geq \lambda_n(H)$ denote the eigenvalues of $H\in Herm(n)$ in decreasing order. We also use the notation $\lambda_{\max}(H) := \lambda_1(H)$, $\lambda_{\min}(H) := \lambda_n(H)$. Furthermore, $E_k(H)$ stands for the eigenspace belonging to $\lambda_k(H)$,

$$E_k(H) := \{ v\in\mathbb{C}^n \mid \; Hv = \lambda_k(H)\, v \}.$$

Note that $\lambda_k(H) = v^*Hv$ for all $v \in E_k(H)$ with $\|v\| = 1$. Let $\mathcal{S}_k$ denote the set of k-dimensional subspaces of $\mathbb{C}^n$. The Courant-Fischer principle states that

$$\lambda_k(H) = \max_{\mathcal{V}\in\mathcal{S}_k} \min_{\substack{v \in \mathcal{V} \\ \|v\| = 1}} v^*Hv \;=\; \min_{\mathcal{V}\in\mathcal{S}_{n-k+1}} \max_{\substack{v \in \mathcal{V} \\ \|v\| = 1}} v^*Hv. \tag{3.1}$$

In particular,

$$\lambda_{\max}(H) = \max_{\substack{v \in \mathbb{C}^n \\ \|v\| = 1}} v^*Hv, \qquad \lambda_{\min}(H) = \min_{\substack{v \in \mathbb{C}^n \\ \|v\| = 1}} v^*Hv. \tag{3.2}$$

Furthermore, (3.1) implies the following well known inclusion result for the eigenvalues of $H + F$.

$$\lambda_k(H) + \lambda_{\min}(F) \leq \lambda_k(H + F) \leq \lambda_k(H) + \lambda_{\max}(F), \qquad H, F \in Herm(n). \tag{3.3}$$

Let $\mathcal{I} \subseteq \mathbb{R}$ be an interval and let $H : \mathcal{I} \to Herm(n)$ be an analytic function. We consider the maps

$$\phi_k : \mathcal{I} \to \mathbb{R} \qquad \phi_k(t) = \lambda_k(H(t)).$$

Suppose that the eigenvalue $\phi_k(\tau)$ has constant multiplicity for all τ in a neighbourhood of $t \in \mathcal{I}$. It is well known that in this case ϕ_k is differentiable at t and the derivative satisfies

$$\dot{\phi}_k(t) = v^* \dot{H}(t) v,$$

where $\dot{H}(t)$ denotes the derivative of $H(\cdot)$ at t and $v \in E_k(H(t))$, $\|v\| = 1$. If ϕ_k changes multiplicity at t then the function $\phi_k(\cdot)$ is not necessarily differentiable at t. However, the right and the left derivative always exist.

Proposition 3.1. *Let $t \in \mathcal{I}$. Suppose that $\phi_k(t) = \phi_j(t)$ for $k_0 \le j \le k_1$ and $\phi_k(t) \ne \phi_j(t)$ for $1 \le j < k_0$ and $k_1 < j \le n$. Let $V \in \mathbb{C}^{n\times(k_1-k_0+1)}$ be a matrix whose columns form an orthonormal basis of the eigenspace $E_k(H(t))$. The right derivative of $\phi_k(\cdot)$ at t exists and is given by*

$$\lim_{h\to 0+} \frac{\phi_k(t+h) - \phi_k(t)}{h} = \lambda_{k_1-k+1}(V^* \dot{H}(t) V) = v^* \dot{H}(t) v,$$

where $v \in E_k(H(t))$ is any unit vector satisfying $v = V\xi$ for some $\xi \in E_{k_1-k+1}(V^*\dot{H}(t)V)$, $\|\xi\| = 1$. The left derivative of $\phi_k(\cdot)$ at t is given by

$$\lim_{h\to 0+} \frac{\phi_k(t) - \phi_k(t-h)}{h} = \lambda_{k-k_0+1}(V^* \dot{H}(t) V) = v^* \dot{H}(t) v,$$

where $v \in E_k(H(t))$ is any unit vector satisfying $v = V\xi$ for some $\xi \in E_{k-k_0+1}(V^*\dot{H}(t)V)$, $\|\xi\| = 1$.

Proof. This follows from [2, page 149]. See also [19, Theorem 8.4]. □

Corollary 3.2. *Suppose t is not the right boundary point of $\mathcal{I}$ and the function $\phi_k : \mathcal{I} \to \mathbb{R}$ attains a local minimum at t. Assume further that $\phi_k(t) > \phi_{k+1}(t)$. Then $v^*\dot{H}(t)v \ge 0$ for all $v \in E_k(H(t))$.*

The statement below can be found in [12]. An analogous result for singular values has been derived in [21, 23]. We give a proof for completeness.

Proposition 3.3. *Let t be an interior point of $\mathcal{I}$ and suppose that $\phi_k(\cdot)$ attains a local extremum at t. Then there exists $0 \ne v \in E_k(H(t))$ such that $v^*\dot{H}(t)v = 0$.*

Proof. Assume the local extremum is a minimum. Then the right derivative of ϕ_k is nonnegative and the left derivative is nonpositive. Hence, according to Proposition 3.1 there exist $v_0, v_1 \in E_k(H(t)) \setminus \{0\}$ such that $v_0^*\dot{H}(t)v_0 \ge 0$ and $v_1^*\dot{H}(t)v_1 \le 0$. If these inequalities are strict in both cases then v_0 and v_1 are linearly independent, and hence the vectors $v_\theta = (1-\theta)\,v_0 + \theta\, v_1 \in E_k(H(t))$ are nonzero for all $\theta \in [0,1]$. By continuity we have $v_\theta^*\dot{H}(t)v_\theta = 0$ for some θ. The proof for a local maximum is analogous. □

4. Computation of $\underline{m}_h(H_0,H_1)$ and $\overline{m}_h(H_0,H_1)$. In this section we provide useful characterizations of the quantities $\underline{m}_h(H_0,H_1)$, $\overline{m}_h(H_0,H_1)$ defined in (2.8).

Theorem 4.1. *Let $H_0, H_1 \in Herm(n)$ and $\phi(t) = \lambda_{\min}(H_0 + t\,H_1)$, $t \in \mathbb{R}$.*

(i) The function $t \mapsto \phi(t)$ is quasiconcave [1], and

$$\underline{m}_h(H_0,H_1) = \sup_{t\in\mathbb{R}} \phi(t). \tag{4.1}$$

(ii) If H_1 is indefinite then the supremum in (4.1) is attained at some t_0 in the interval $[t_1,t_2]$, where

$$t_1 = -\frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\max}(H_1)}, \qquad t_2 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{|\lambda_{\min}(H_1)|}. \tag{4.2}$$

(iii) If H_1 is semidefinite but not definite then

$$\underline{m}_h(H_0,H_1) = \lambda_{\min}(V^* H_0 V), \tag{4.3}$$

where V is a matrix whose columns form an orthonormal basis of $\ker H_1$.

[1] Let $\mathcal{I}$ be an interval. A function $f : \mathcal{I} \to \mathbb{R}$ is said to be quasiconcave if each superlevel set $\{x \in \mathcal{I};\ f(x) \ge c\}$, $c \in \mathbb{R}$, is an interval. $f : \mathcal{I} \to \mathbb{R}$ is said to be quasiconvex if each sublevel set $\{x \in \mathcal{I};\ f(x) \le c\}$, $c \in \mathbb{R}$, is an interval. A continuous function $f : \mathcal{I} \to \mathbb{R}$ is quasiconcave (quasiconvex) iff each local extremum of f is a global maximum (minimum).

(iv) If H_1 is definite then $\underline{m}_h(H_0,H_1)=\infty$.
(v) If H_1 is positive (negative) semidefinite then the function $\phi(\cdot)$ is increasing (decreasing).
(vi) If H_1 is positive (negative) definite then the function $\phi(\cdot)$ is strictly increasing (strictly decreasing) and $\lim_{t\to\infty}\phi(t)=\infty$ ($\lim_{t\to-\infty}\phi(t)=\infty$).

Proof. We will need the following inequalities which are consequences of (3.3).

$$\lambda_{\min}(H_0)+\lambda_{\min}(t\,H_1)\le\phi(t)\le\lambda_{\max}(H_0)+\lambda_{\min}(t\,H_1), \tag{4.4}$$

$$\phi(t_*)+\lambda_{\min}(t\,H_1)\le\phi(t_*+t)\le\phi(t_*)+\lambda_{\max}(t\,H_1), \qquad t,t_*\in\mathbb{R}. \tag{4.5}$$

Note that

$$\lambda_{\min}(t\,H_1)=\begin{cases}\lambda_{\min}(H_1)\,t & \text{if } t\ge 0,\\ \lambda_{\max}(H_1)\,t & \text{if } t\le 0.\end{cases}$$

For any unit vector $v\in\mathbb{C}^n$ satisfying $v^*H_1v=0$ and any $t\in\mathbb{R}$ we have by (3.2),

$$\phi(t)\le v^*(H_0+t\,H_1)v=v^*H_0v.$$

This implies

$$\sup_{t\in\mathbb{R}}\phi(t)\le\underline{m}_h(H_0,H_1). \tag{4.6}$$

Suppose now that the function $\phi(\cdot)$ attains a local extremum at t_0. Then by Proposition 3.3 there exists a unit vector v_0 satisfying $(H_0+t_0\,H_1)v_0=\phi(t_0)\,v_0$ and $v_0^*H_1v_0=0$, whence $v_0^*H_0v_0=\phi(t_0)$. Thus $\underline{m}_h(H_0,H_1)\le\phi(t_0)\le\sup_{t\in\mathbb{R}}\phi(t)$, and then (4.6) yields

$$\sup_{t\in\mathbb{R}}\phi(t)=\phi(t_0)=\underline{m}_h(H_0,H_1).$$

Thus each local extremum of ϕ is the global maximum. This implies that the function ϕ is quasiconcave. Suppose now, that H_1 is indefinite and let t_1,t_2 be defined as in (4.2). Then for $t\notin[t_1,t_2]$, $\lambda_{\max}(H_0)+\lambda_{\min}(t\,H_1)\le\lambda_{\min}(H_0)=\phi(0)$. By combining this with (4.4) we obtain $\phi(t)\le\phi(0)$. Consequently, ϕ attains a local maximum at some $t_0\in[t_1,t_2]$. We thus have shown *(ii)* and (4.1) for the case that H_1 is indefinite. Suppose now, that H_1 is semidefinite. Then $v^*H_1v=0$ if and only if $v\in\ker H_1$. If H_1 is definite then $\ker H_1=\{0\}$. This implies *(iii)* and *(iv)*. *(v)* and *(vi)* are immediate consequences of (4.4) and (4.5). Thus (4.1) holds for the case that H_1 is definite. We now prove (4.1) for the semidefinite case by contradiction. Assume without loss of generality that H_1 is positive semidefinite but not definite. Assume further, that (4.1) fails. Then by *(v)* and (4.6) there are unit vectors v_k, $k\in\mathbb{N}$, and an $\epsilon>0$ such that

$$v_k^*(H_0+k\,H_1)v_k=\phi(k)<\underline{m}_h(H_0,H_1)-\epsilon,$$

whence

$$0\le v_k^*H_1v_k\le\frac{1}{k}(\underline{m}_h(H_0,H_1)-v_k^*H_0v_k-\epsilon). \tag{4.7}$$

By compactness there is a subsequence v_{k_j} which converges to a unit vector $\hat v$. It follows from (4.7) that $\hat v^*H_1\hat v=0$. But then $\hat v^*H_0\hat v\ge\underline{m}_h(H_0,H_1)$ by definition of $\underline{m}_h(H_0,H_1)$. Therefore $\underline{m}_h(H_0,H_1)-v_{k_j}^*H_0v_{k_j}-\epsilon<0$ for k_j large enough. The latter contradicts (4.7). ☐

REMARK 4.2. *The function*

$$(H_0,H_1)\mapsto\underline{m}_h(H_0,H_1) \tag{4.8}$$

is discontinuous at (H_0,H_1) if H_1 is semidefinite but not definite, since then an arbitrarily small perturbation of H_1 can change the dimension of $\ker H_1$. However, if H_1 is indefinite then the function (4.8) is continuous at (H_0,H_1). This is seen as follows. If H_1 is indefinite, then H_1+E_1 is indefinite for all $E_1\in\mathcal{U}_\epsilon=\{\,E\in Herm(n);\ \|E\|\le\epsilon\}$, $\epsilon>0$ sufficiently small. Let

$$\begin{aligned}a&=\max\{\,\lambda_{\max}(H_0+E)-\lambda_{\min}(H_0+E);\ \ E\in\mathcal{U}_\epsilon\,\},\\ T_1&=-a/\min\{\,\lambda_{\max}(H_1+E);\ \ E\in\mathcal{U}_\epsilon\,\},\\ T_2&=a/\min\{\,|\lambda_{\min}(H_1+E)|;\ \ E\in\mathcal{U}_\epsilon\,\}.\end{aligned}$$

Then by part (ii) of Theorem 4.1 we have for all $(E_0, E_1) \in \mathcal{U}_\epsilon^2$,

$$\underline{m}_h(H_0 + E_0, H_1 + E_1) = \max_{t \in [T_1, T_2]} \lambda_{\min}(H_0 + E_0 + t(H_1 + E_1)).$$

The right hand side of the latter identity is a continuous function of (E_0, E_1) *since the maximum is taken over a compact set.*

REMARK 4.3. *The quasiconcavity of the function* $t \mapsto \lambda_{\min}(H_0 + tH_1), t \in \mathbb{R}$, *can also be shown in the following way. The first relation in (3.2) yields that the function* $t \mapsto \lambda_{\max}(H_0 + tH_1), t \in \mathbb{R}$, *is convex for any* $H_0, H_1 \in Herm(n)$. *Thus* $t \mapsto -\lambda_{\max}((-H_0) + t(-H_1)) = \lambda_{\min}(H_0 + tH_1)$ *is quasiconcave.*

REMARK 4.4. *If* H_1 *is indefinite the* $\underline{m}_h(H_0, H_1)$ *can easily be computed via (4.1) and (4.2). If* H_1 *is semidefinite but not definite then (4.3) can be used for computation.*

For $\overline{m}_h(H_0, H_1)$ we have the following result. Its proof is analogous to that of Theorem 4.1. and therefore omitted.

THEOREM 4.5. *Let* $H_0, H_1 \in Herm(n)$ *and* $\phi(t) = \lambda_{\max}(H_0 + t\,H_1)$, $t \in \mathbb{R}$.

(i) The function $t \mapsto \phi(t)$ *is convex , and*

$$\overline{m}_h(H_0, H_1) = \inf_{t \in \mathbb{R}} \phi(t). \tag{4.9}$$

(ii) If H_1 *is indefinite then the infimum in (4.9) is attained at some* t_0 *in the interval* $[t_1, t_2]$, *where*

$$t_1 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\min}(H_1)}, \qquad t_2 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\max}(H_1)}. \tag{4.10}$$

(iii) If H_1 *is semidefinite but not definite then*

$$\underline{m}_h(H_0, H_1) = \lambda_{\max}(V^* H_0 V), \tag{4.11}$$

where V *is a matrix whose columns form an orthonormal basis of* $\ker H_1$.

(iv) If H_1 *is definite then* $\overline{m}_h(H_0, H_1) = -\infty$.

(v) If H_1 *ist positive (negative) semidefinite then the function* $\phi(\cdot)$ *is increasing (decreasing).*

(vi) If H_1 *is positive (negative) definite then the function* $\phi(\cdot)$ *is strictly increasing (strictly decreasing) and* $\lim_{t \to -\infty} \phi(t) = -\infty$ $(\lim_{t \to \infty} \phi(t) = -\infty)$.

5. Computation of $\underline{m}_{hs}(H, S)$ and $\overline{m}_{hs}(H, S)$. We now treat the Hermitian-symmetric case. Recall the definition in (2.8),

$$\overline{m}_{hs}(H, S) = \sup\{\, v^* H v \,;\ v \in \mathbb{C}^n,\ v^\top S v = 0,\ \|v\| = 1 \,\}, \qquad H \in Herm(n),\ S \in Sym(n).$$

If $S = 0$ then $\overline{m}_{hs}(H, S) = \lambda_{\max}(H)$. Suppose that $\operatorname{rank}(S) = 1$. Then S can be written in the form $S = xx^\top$ for some nonzero $x \in \mathbb{C}^n$. Hence, for any $v \in \mathbb{C}^n$, $v^\top S v = (v^\top x)^2$. Let $\mathcal{V} = \{\, v \in \mathbb{C}^n \,;\ v^\top x = 0 \,\}$. Then

$$\overline{m}_{hs}(H, S) = \overline{m}_{hs}(H, xx^\top) = \sup_{v \in \mathcal{V}, \|v\| = 1} v^* H v. \tag{5.1}$$

If $n = 1$ then $\mathcal{V} = 0$. Thus $\overline{m}_{hs}(H, xx^\top) = -\infty$. Suppose $n \geq 2$ and let $V \in \mathbb{C}^{n \times (n-1)}$ be a matrix whose columns form an orthonormal basis of $\mathcal{V}$. Then (5.1) yields

$$\overline{m}_{hs}(H, xx^\top) = \max_{\xi \in \mathbb{C}^{n-1}, \|\xi\| = 1} (V\xi)^* H (V\xi) = \lambda_{\max}(V^* H V).$$

If $\operatorname{rank}(S) \geq 2$ then $\overline{m}_{hs}(H, S)$ can be computed by minimizing the second largest eigenvalue of a hermitian pencil. The precise statement is as follows.

THEOREM 5.1. *Suppose* $\operatorname{rank}(S) \geq 2$. *Let* $t_1 = 2\|H\|/\sigma_2(S)$, *where* $\sigma_2(S)$ *denotes the second largest singular value of* S. *Then*

$$\overline{m}_{hs}(H, S) = \min_{0 \leq t \leq t_1} \lambda_2 \left(\begin{bmatrix} H & t\overline{S} \\ tS & \overline{H} \end{bmatrix} \right).$$

The function to be minimized is quasiconvex.

We split the proof into several lemmas. Let us introduce some notation.

- For $t \in \mathbb{R}$ we set

$$F(t) := \begin{bmatrix} H & t\,\overline{S} \\ tS & \overline{H} \end{bmatrix}, \qquad \phi(t) := \lambda_2(F(t)).$$

- For a unit vector $v \in \mathbb{C}^n$,

$$\mathcal{U}_v := \left\{ \begin{bmatrix} z_1\, v \\ \overline{z_2\, v} \end{bmatrix};\ z_1, z_2 \in \mathbb{C} \right\}.$$

Note that $\mathcal{U}_v$ is a 2-dimensional subspace of $\mathbb{C}^{2n}$, and

$$\begin{bmatrix} z_1\, v \\ \overline{z_2\, v} \end{bmatrix}^* F(t) \begin{bmatrix} z_1\, v \\ \overline{z_2\, v} \end{bmatrix} = (|z_1|^2 + |z_2|^2)\, v^* H v + 2t\, \Re(z_1 z_2\, v^\top S v), \quad z_1, z_2 \in \mathbb{C}. \tag{5.2}$$

Lemma 5.1. *For any $H \in Herm(n)$, $S \in Sym(n)$ we have $\overline{m}_{hs}(H,S) \le \inf_{t\in\mathbb{R}} \phi(t)$.*

Proof. Let $v \in \mathbb{C}^n$ be a unit vector satisfying $v^\top S v = 0$. Thus by the Courant-Fischer max-min-principle and (5.2),

$$\phi(t) \ge \min_{x \in \mathcal{U}_v,\, \|x\|=1} x^* H x = v^* H v \quad \text{ for all } t \in \mathbb{R}.$$

Hence, $\phi(t) \ge \overline{m}_{hs}(H,S)$. □

Lemma 5.2. *Let $v_1, \ldots, v_d$ be a basis of the eigenspace $E_k(H)$. Then*

$$E_{2k-1}(F(0)) = E_{2k}(F(0)) = \bigoplus_{j=1}^d \mathcal{U}_{v_j},$$

where $\oplus$ denotes the direct sum.

The simple proof is left to the reader.

Lemma 5.3. *Let $\mathcal{V}$ be a subspace of $\mathbb{C}^n$ of dimension $\dim \mathcal{V} \ge 2$. Then to any $S \in Sym(n)$ there is a nonzero $v \in \mathcal{V}$ satisfying $v^\top S v = 0$.*

Proof. For $z_1, z_2 \in \mathbb{C}$ let $v_{z_1,z_2} = z_1\, v_1 + z_2\, v_2$, where $v_1, v_2 \in \mathcal{V}$ are linearly independent vectors. The function $(z_1, z_2) \mapsto v_{z_1,z_2}^\top S\, v_{z_1,z_2}$ is a homogeneous quadratic polynomial and has a zero $(z_1, z_2) \ne (0,0)$. □

Lemma 5.4. *The following statements are equivalent.*

(i) $\overline{m}_{hs}(H,S) = \phi(0) = \lambda_{\max}(H)$.

(ii) Either $\dim E_1(H) \ge 2$, *or* $\dim E_1(H) = 1$ *and $v^\top S v = 0$ for $v \in E_1(H)$.*

(iii) The function $\mathbb{R} \ni t \mapsto \phi(t)$ attains its minimum at $t = 0$.

Proof. By Lemma 5.2 we have $\lambda_{\max}(H) = \phi(0)$.

$(i) \Leftrightarrow (ii)$. First note that $\overline{m}_{hs}(H,S) \le \max\{\, v^* H v;\ v \in \mathbb{C}^n,\ \|v\| = 1\} = \lambda_{\max}(H)$. Equality holds if and only if there is a unit vector $v \in E_1(H)$ such that $v^\top S v = 0$. By Lemma 5.3 the latter condition is satisfied if $\dim E_1(H) \ge 2$.

The implication $(i) \Rightarrow (iii)$ follows from Lemma 5.1.

$(iii) \Rightarrow (i)$. Since (i) is satisfied if $\dim E_1(H) \ge 2$ we may assume that $\dim E_1(H) = 1$. Then by Lemma 5.2, $E_2(F(0)) = E_1(F(0)) = \mathcal{U}_v$ for a unit vector $v \in E_1(H)$. Furthermore, from (iii) and Corollary 3.2 it follows that $0 \le x^* \dot{F}(0) x$ for all $x \in E_2(F(0))$. In other words, we have for all $z_1, z_2 \in \mathbb{C}$,

$$0 \le \begin{bmatrix} z_1\, v \\ \overline{z_2\, v} \end{bmatrix}^* \begin{bmatrix} 0 & \overline{S} \\ S & 0 \end{bmatrix} \begin{bmatrix} z_1\, v \\ \overline{z_2\, v} \end{bmatrix} = 2\, \Re(z_1\, z_2\, v^\top S v).$$

This implies $v^\top S v = 0$. Thus $\overline{m}_{hs}(H,S) = v^* H v = \phi(0)$. □

Lemma 5.5. *Suppose the function $\mathbb{R} \ni t \mapsto \phi(t)$ has a local extremum at $t_0 \ne 0$. Then there is a unit vector $v \in \mathbb{C}^n$ satisfying $v^* H v = \phi(t_0)$ and $v^\top S v = 0$. Hence, $\phi(t_0) \le \overline{m}_{hs}(H,S)$.*

Proof. If the assumption of the Lemma holds then by Proposition 3.3 there is a nonzero $v_0 \in \mathbb{C}^{2n}$, such that

$$F(t_0)\, v_0 = \phi(t_0)\, v_0, \tag{5.3}$$

$$v_0^* \dot{F}(t_0) v_0 = 0. \tag{5.4}$$

Let

$$H_0 := H - \phi(t_0)\, I_n, \qquad v_0 = \begin{bmatrix} x \\ \overline{y} \end{bmatrix}, \quad x, y \in \mathbb{C}^n.$$

Then (5.3) is equivalent to the equations

$$H_0 x = -t_0 \overline{Sy}, \qquad \overline{H_0 y} = -t_0\, Sx, \tag{5.5}$$

which imply

$$\begin{aligned} x^* H_0 x &= -t_0\, \overline{x^\top S y} = y^* H_0 y \\ x^* H_0 y &= -t_0\, \overline{x^\top S x} = -t_0\, y^\top S y. \end{aligned} \tag{5.6}$$

Since $t_0 \neq 0$ it follows that

$$x^\top S y \in \mathbb{R}, \tag{5.7}$$

$$y^\top S y = \overline{x^\top S x}. \tag{5.8}$$

We have that $v_0^* \dot{F}(t_0) v_0 = 2\Re(x^\top S\, y)$. Thus, (5.4) and (5.7) yield

$$x^\top S y = 0. \tag{5.9}$$

Now let

$$\beta := \begin{cases} 1 & \text{if } x^T S x = 0, \\ i \frac{x^T S x}{|x^T S x|} & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} (x \pm \beta y)^\top S (x \pm \beta y) &= x^T S x + \beta^2 y^\top S y + 2\beta x^\top S y \\ &= x^\top S x + \underbrace{\beta^2 \overline{x^\top S x}}_{= -x^T S x} + 2\beta \underbrace{x^\top S y}_{=0} \qquad \text{(using (5.8) and (5.9))} \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} (x \pm \beta y)^* H_0 (x \pm \beta y) &= x^* H_0 x + |\beta|^2 y^* H_0 y \pm 2\, \Re(x^* H_0 y\, \beta) \\ &= -t_0 \left(\, (1 + |\beta|^2)\, \underbrace{\overline{x^\top S y}}_{=0} \pm 2\, \underbrace{\Re(\overline{x^\top S x}\, \beta)}_{=0} \,\right) \qquad \text{(using (5.6) and (5.9))} \\ &= 0. \end{aligned}$$

At least one of the vectors $x \pm \beta\, y$ is nonzero and can therefore be divided by its norm. The resulting vector $v \in \mathbb{C}^n$ has the required properties. □

LEMMA 5.6. *The function $\mathbb{R} \ni t \mapsto \phi(t)$ satisfies $\phi(t) = \phi(-t)$ for all $t \in \mathbb{R}$. If* $\operatorname{rank}(S) \geq 2$ *then ϕ attains its minimum in the interval $[0, t_1]$, where $t_1 = 2\|H\|/\sigma_2(S)$.*

Proof. Let $T = \begin{bmatrix} -I_n & 0 \\ 0 & I_n \end{bmatrix}$. Then $F(-t) = T\, F(t)\, T^{-1}$. Thus $\phi(t) = \phi(-t)$. Next, we give a lower bound for $\phi(t)$.

The eigenvalues of $\begin{bmatrix} 0 & t\,\overline{S} \\ t\, S & 0 \end{bmatrix} = \begin{bmatrix} 0 & t\, S^* \\ t\, S & 0 \end{bmatrix}$ are the singular values of S and their negatives. In particular,

$$\lambda_2 \left(\begin{bmatrix} 0 & t\,\overline{S} \\ tS & 0 \end{bmatrix} \right) = \sigma_2(tS) = |t|\, \sigma_2(S).$$

Furthermore,

$$\phi(0) = \lambda_2 \left(\begin{bmatrix} H & 0 \\ 0 & \overline{H} \end{bmatrix} \right) = \lambda_{\max} \left(\begin{bmatrix} H & 0 \\ 0 & \overline{H} \end{bmatrix} \right) = \lambda_{\max}(H) \leq \|H\|.$$

We conclude that

$$\phi(t) = \lambda_2\left(\begin{bmatrix} H & t\overline{S} \\ tS & \overline{H} \end{bmatrix}\right) \geq \lambda_2\left(\begin{bmatrix} 0 & t\overline{S} \\ tS & 0 \end{bmatrix}\right) - \lambda_{\max}\left(\begin{bmatrix} H & 0 \\ 0 & \overline{H} \end{bmatrix}\right) \geq |t|\sigma_2(S) - \|H\|.$$

Thus, if $|t| \geq t_1$ then $\phi(t) \geq \phi(0)$. Consequently ϕ attains its minimum at some $t_0 \in \mathbb{R}$ with $|t_0| \leq t_1$. Since $\phi(t) = \phi(-t)$ there exists a minimizer $t_0 \geq 0$. □

We now summarize the results and finish the proof of Theorem 5.1. By Lemma 5.1,

$$\overline{m}_{hs}(H,S) \leq \inf_{t\in\mathbb{R}} \phi(t). \tag{5.10}$$

By Lemma 5.6 the infimum is attained for some $t_0 \in [0, t_1]$. If $t_0 = 0$ is a minimizer then equality holds in (5.10) by Lemma 5.4. If $t_0 \neq 0$ is a local extremum then equality hold in (5.10) by Lemma 5.5. This in particular shows that each local extremum is a global minimum. Thus ϕ is quasiconvex, and the proof is complete.

From Theorem 5.1 and the fact that $\overline{m}_{hs}(H,S) = -\underline{m}_{hs}(-H,-S)$ one easily obtains the following result.

THEOREM 5.2. *Let $H \in Herm(n)$, $S \in Sym(n)$ with* $\operatorname{rank}(S) \geq 2$, *and $t_1 = 2\|H\|/\sigma_2(S)$. Then*

$$\overline{m}_{hs}(H,S) = \max_{0\leq t\leq t_1} \lambda_{2n-1}\left(\begin{bmatrix} H & t\,\overline{S} \\ tS & \overline{H} \end{bmatrix}\right).$$

The function to be maximized is quasiconcave.

REMARK 5.3. *With the same reasoning as in Remark 4.2 one can show that the functions*

$$Herm(n) \times Sym(n) \ni (H,S) \longmapsto \overline{m}_{hs}(H,S),$$
$$Herm(n) \times Sym(n) \ni (H,S) \longmapsto \underline{m}_{hs}(H,S)$$

are continuous at all (H,S) with $\operatorname{rank}(S) \geq 2$.

6. Self- and skew-adjoint matrices. We now treat μ-values with respect to linear subspaces which are induced by a scalar product on $\mathbb{C}^n$. Specifically we show that these μ-values are closely related to the μ-values with respect to Hermitian, symmetric and skew-symmetric perturbations.
For nonsingular $\Pi \in \mathbb{C}^{n\times n}$ we consider the scalar products

$$\langle x, y\rangle_\Pi = x^\star \Pi y, \qquad x, y \in \mathbb{C}^n,\ \star \in \{*, \top\}.$$

Depending on whether $\star = \top$ or $\star = *$ the scalar product is a bilinear form or a sesquilinear form. We assume that Π satisfies a symmetry relation of the form

$$\Pi^\star = \epsilon_0 \Pi, \text{ with } \epsilon_0 = -1 \text{ or } \epsilon_0 = 1. \tag{6.1}$$

A matrix $\Delta \in \mathbb{C}^{n\times n}$ is said to be self-adjoint (skew-adjoint) with respect to the scalar product $\langle \cdot, \cdot \rangle_\Pi$ if

$$\langle \Delta x, y\rangle_\Pi = \epsilon\, \langle x, \Delta y\rangle_\Pi \quad \text{for all } x, y \in \mathbb{C}^n, \tag{6.2}$$

and $\epsilon = 1$ ($\epsilon = -1$). The relation (6.2) is easily seen to be equivalent to

$$\Delta^\star \Pi = \epsilon\, \Pi \Delta. \tag{6.3}$$

We denote the sets of self- and skew-adjoint matrices by

$$struct(\Pi, \star, \epsilon) := \{\ \Delta \in \mathbb{C}^{n\times n}\ ;\ \ \Delta^\star \Pi = \epsilon\, \Pi \Delta\ \}.$$

The relation (6.1) implies that (6.3) is equivalent to

$$(\Pi\Delta)^\star = \epsilon_0\, \epsilon\, \Pi\Delta. \tag{6.4}$$

We thus have the lemma below.

Lemma 6.1. *Let $\Pi, \Delta \in \mathbb{C}^{n\times n}$. Suppose $\Pi^\star = \epsilon_0 \Pi$ with $\epsilon_0 = -1$ or $\epsilon_0 = 1$. Then the following equivalences hold.*

$$\Delta \in struct(\Pi, \star, \epsilon) \Leftrightarrow \begin{cases} \Pi\Delta \in Herm(n) & \text{if } \epsilon_0\epsilon = 1,\ \star = *, \\ \Pi\Delta \in Sym(n) & \text{if } \epsilon_0\epsilon = 1,\ \star = \top, \\ \Pi\Delta \in Skew(n) & \text{if } \epsilon_0\epsilon = -1,\ \star = \top, \\ \pm i\,\Pi\Delta \in Herm(n) & \text{if } \epsilon_0\epsilon = -1,\ \star = *. \end{cases}$$

In many applications Π is unitary. The most common examples are $\Pi \in \{\mathrm{diag}(I_k, -I_{n-k}), E_n, J_n\}$, where

$$J_n := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{C}^{2n\times 2n}, \qquad E_n := \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{6.5}$$

For unitary Π the μ-values of the associated self- and skew-adjoint classes can be expressed in terms of the μ-values for $Herm(n)$, $Sym(n)$ and $Skew(n)$:

Corollary 6.1. *Suppose $\Pi \in \mathbb{C}^{n\times n}$ is unitary and satisfies $\Pi^\star = \epsilon_0\Pi$ with $\epsilon_0 = -1$ or $\epsilon_0 = 1$. Let $struct = struct(\Pi, \star, \epsilon)$. Then for any $M \in \mathbb{C}^{n\times n}$,*

$$\mu_{struct}(M) = \begin{cases} \mu_{Herm}(M\Pi^*) & \text{if } \epsilon_0\epsilon = 1,\ \star = *, \\ \mu_{Sym}(M\Pi^*) & \text{if } \epsilon_0\epsilon = 1,\ \star = \top, \\ \mu_{Skew}(M\Pi^*) & \text{if } \epsilon_0\epsilon = -1,\ \star = \top, \\ \mu_{Herm}(\pm i\, M\Pi^*) & \text{if } \epsilon_0\epsilon = -1,\ \star = *, \end{cases} \tag{6.6}$$

and

$$\widetilde{\mu}_{struct}(M) = \begin{cases} \widetilde{\mu}_{Herm}(\Pi M) & \text{if } \epsilon_0\epsilon = 1,\ \star = *, \\ \widetilde{\mu}_{Sym}(\Pi M) & \text{if } \epsilon_0\epsilon = 1,\ \star = \top, \\ \widetilde{\mu}_{Skew}(\Pi M) & \text{if } \epsilon_0\epsilon = -1,\ \star = \top, \\ \widetilde{\mu}_{Herm}(\pm i\, \Pi M) & \text{if } \epsilon_0\epsilon = -1,\ \star = *. \end{cases} \tag{6.7}$$

Proof. Since Π is unitary, we have

$$\begin{aligned} \mu_{struct}(M) &= (\ \inf\{\ \|\Delta\|\ ;\ \Delta \in struct, \quad 1 \in \sigma(\Delta\, M)\ \}\ )^{-1} \\ &= (\ \inf\{\ \|\Pi\Delta\|\ ;\ \Delta \in struct, \quad 1 \in \sigma((\Pi\Delta)\,(M\Pi^*))\ \}\ )^{-1}, \end{aligned} \tag{6.8}$$

$$\begin{aligned} \widetilde{\mu}_{struct}(M) &= \inf\{\ \|\Delta\|\ ;\ \Delta \in struct, \quad \det(M - \Delta) = 0\ \} \\ &= \inf\{\ \|\Pi\Delta\|\ ;\ \Delta \in struct, \quad \det(\Pi M - \Pi\Delta) = 0\ \}. \end{aligned} \tag{6.9}$$

Thus, the first three identities in (6.6) and (6.7) are consequences of the first three equivalences in Lemma 6.1. On replacing in (6.8) and (6.9) Π by $\pm i\,\Pi$ one obtains the fourth identity in (6.6) and (6.7) from the fourth equivalence in Lemma 6.1. ◻

7. Application: Spectral value sets for Hamiltonian matrices. A matrix which is skew-adjoint with respect to the sesquilinear form induced by J_n is called Hamiltonian. Let

$$Ham(n) := \{\ \Delta \in \mathbb{C}^{2n\times 2n}\ ;\ \Delta^* J_n = -J_n\Delta\ \}$$

denote the set of complex Hamiltonian matrices. Each $H \in Ham(n)$ has block structure

$$H = \begin{bmatrix} A & C \\ B & -A^* \end{bmatrix} \qquad \text{with } A \in \mathbb{C}^{n\times n} \text{ and } B, C \in Herm(n).$$

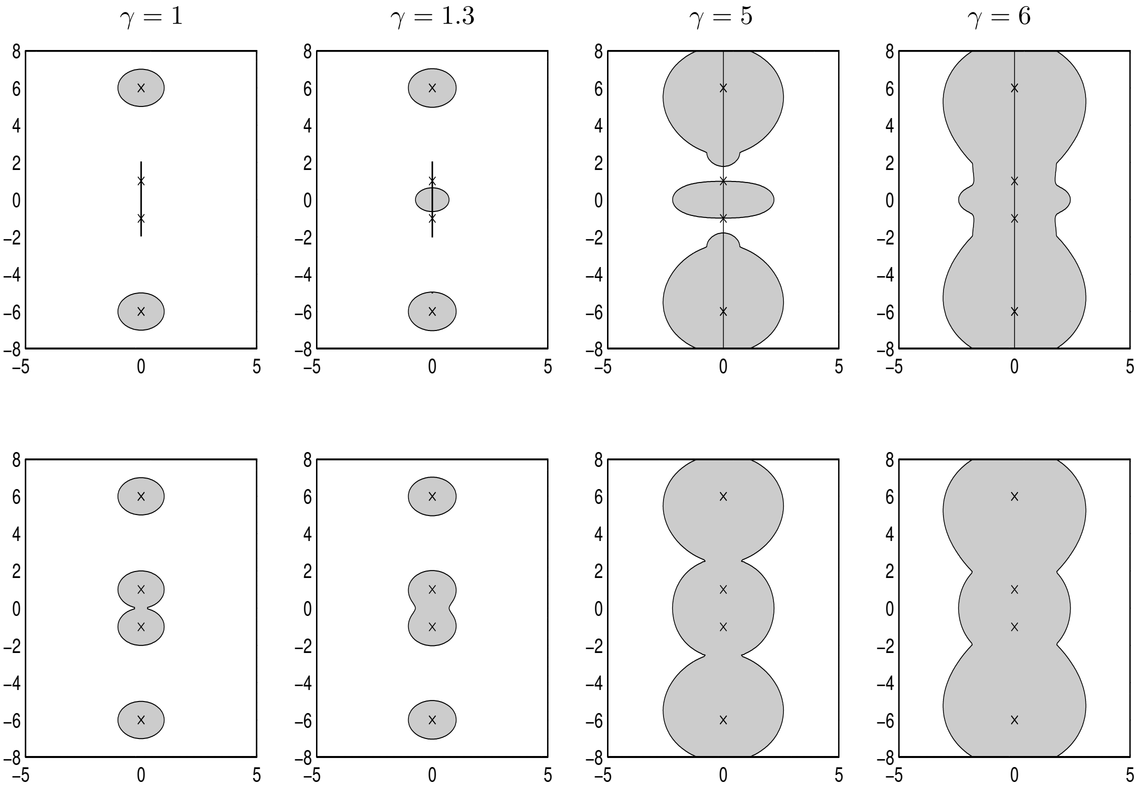

FIG. 7.1. *The sets $\sigma_{Ham}(H_\gamma;1)$ (upper row) and $\sigma_{\mathbb{C}^{n\times n}}(H_\gamma;1)$ (lower row).*

The spectral value sets of H with respect to unstructured perturbations are by (1.3) and (1.11),

$$\sigma_{\mathbb{C}^{n\times n}}(H;\delta) = \bigcup_{\Delta\in\mathbb{C}^{n\times n},\ \|\Delta\|<\delta} \sigma(H+\Delta) \ = \ \{s\in\mathbb{C};\ \sigma_{\min}(s\,I_n - H) < \delta\,\}, \quad \delta>0.$$

The spectral value sets with respect to Hamiltonian perturbations are

$$\sigma_{Ham}(H;\delta) = \bigcup_{\Delta\in Ham(n),\ \|\Delta\|<\delta} \sigma(H+\Delta) \ = \ \{s\in\mathbb{C};\ f(s) < \delta\,\}, \quad \delta>0, \tag{7.1}$$

where

$$f:\mathbb{C}\to\mathbb{R}, \qquad f(s) := \widetilde{\mu}_{Ham}(s\,I_n - H). \tag{7.2}$$

By Corollary 6.1 we have $f(s) = \widetilde{\mu}_{Herm}(\Phi(s))$, where $\Phi(s) = J_n(sI_n - H) = \begin{bmatrix} -B & sI_n + A^* \\ -sI_n + A & C \end{bmatrix}$. Since $\frac{\Phi(s)-\Phi(s)^*}{2i} = -i\,(\Re s)\,J_n$ and $\Phi(s)^*\Phi(s) = (s\,I_n - H)^*(s\,I_n - H)$ we obtain from Corollary 2.3 and Theorem 4.1,

$$\begin{aligned}
f(s) &= \sqrt{\underline{m}_h(\,\Phi(s)^*\Phi(s),\ \tfrac{\Phi(s)-\Phi(s)^*}{2i}\,)} \\
&= \sqrt{\underline{m}_h(\,(s\,I_n - H)^*(s\,I_n - H),\ -i\,(\Re s)\,J_n\,)} \\
&= \begin{cases} \sqrt{\underline{m}_h((s\,I_n - H)^*(s\,I_n - H),\,0)} & \text{if } s\in i\mathbb{R}, \\ \sqrt{\underline{m}_h((s\,I_n - H)^*(s\,I_n - H),\,i\,J_n)} & \text{otherwise,} \end{cases} \\
&= \begin{cases} \sigma_{\min}(s\,I_n - H) & \text{if } s\in i\mathbb{R}, \\ \sqrt{\max\limits_{t\in\mathbb{R}} \lambda_{\min}\ ((s\,I_n - H)^*(s\,I_n - H) + t\,i\,J_n)} & \text{otherwise.} \end{cases}
\end{aligned} \tag{7.3}$$

Furthermore, by (4.2) the maximum in (7.3) is attained for some $t \in \mathbb{R}$ satisfying

$$|t| \leq \sigma_{\max}(s\,I_n - H)^2 - \sigma_{\min}(s\,I_n - H)^2.$$

REMARK 7.1. *Since $f(s) = \sigma_{\min}(s\,I_n - H)$ for $s \in i\mathbb{R}$, the structured and the unstructured spectral value sets coincide on the imaginary axis. Precisely, $i\mathbb{R} \cap \sigma_{Ham}(H;\delta) = i\mathbb{R} \cap \sigma_{\mathbb{C}^{n\times n}}(H;\delta)$.*

PROPOSITION 7.2. *Let $H \in Ham(n)$, and consider the function f defined in (7.2).*

(i) The restriction of f to the set $\mathbb{C} \setminus i\mathbb{R}$ is continuous. Hence, for each $\delta > 0$ the set $\sigma_{Ham}(H;\delta) \setminus i\mathbb{R}$ is an open subset of $\mathbb{C}$.

(ii) Let $s \in \mathbb{C} \setminus i\mathbb{R}$. Then $f(s) = \sigma_{\min}(s\,I_n - H)$ iff there exists a right singular vector $v \neq 0$ to $\sigma_{\min}(s\,I_n - H)$ satisfying $v^* J_n v = 0$.

(iii) Let $s \in i\mathbb{R}$ be an eigenvalue of H. Let $E = \{\, v \in \mathbb{C}^n;\ Hv = s\,v \,\}$ be the associated eigenspace, and let $D(s,\epsilon) = \{\, z \in \mathbb{C};\ \ |s - z| < \epsilon \,\}$ denote the open disk of radius $\epsilon > 0$ about s. Then the following assertions are equivalent.

(a) The function f is discontinuous at s.

(b) We have $v^* J_n v \neq 0$ for all $v \in E \setminus \{0\}$.

(c) There exist $\epsilon > 0$ and $\delta > 0$ sucht that $\sigma_{Ham}(H;\delta) \cap D(s,\epsilon) \subset i\mathbb{R}$.

Proof. From Remark 4.2 we obtain that the function

$$g : \mathbb{C} \to \mathbb{R}, \qquad g(s) = \sqrt{\underline{m}_h((s\,I_n - H)^*(s\,I_n - H),\, i\,J_n)}$$

is continuous. By definition of $\underline{m}_h$ we have for all $s \in \mathbb{C}$ that $g(s) \leq \sigma_{\min}(s\,I_n - H)$. Equality holds iff there exists a right singular vector $v \neq 0$ to the singular value $\sigma_{\min}(s\,I_n - H)$ such that $v^* J_n v = 0$. If s is an eigenvalue of H then the singular vectors to $\sigma_{\min}(s\,I_n - H) = 0$ are the eigenvectors belonging to s. Hence, all statements of the proposition follow from (7.1) and the fact that $f(s) = g(s)$ for $s \notin i\mathbb{R}$ and $f(s) = \sigma_{\min}(s\,I_n - H)$ for $s \in i\mathbb{R}$. □

The upper row in Figure 7.1 shows the spectral value sets $\sigma_{Ham}(H_\gamma;1)$ for the Hamiltonian matrices

$$H_\gamma = \begin{bmatrix} 0 & \gamma^{-1} C \\ \gamma B & 0 \end{bmatrix}, \quad B = \operatorname{diag}(1,6,-6), \quad C = \operatorname{diag}(1,6,-6), \quad \gamma \in \{1, 1.3, 5, 6\}.$$

The lower row in the figure shows the sets $\sigma_{\mathbb{C}^{n\times n}}(H_\gamma;1)$ for comparison. The crosses mark the eigenvalues of H_γ.

Acknowledgments. The author thanks Daniel Kressner for valuable comments on an earlier version of this paper.

REFERENCES

[1] G. J. Balas, J. C. Doyle, K. Glover, A. Packard, and R. Smith. *μ-Analysis and Synthesis Toolbox. User's Guide.* The MathWorks Inc., South Natick, Massachusetts, 1991.

[2] H. Baumgärtel. *Analytic perturbation theory for matrices and operators. Operator Theory: Advances and Applications. Vol. 15.* Birkhäuser Verlag, Basel - Boston - Stuttgart, 1985

[3] B. Bernhardsson, A. Rantzer and L. Qiu. Real perturbation values and real quadratic forms in a complex vector space. *Linear Algebra Appl.* 270, 131-154 (1998).

[4] N. A. Bruinsma and M. Steinbuch. A fast algorithm to compute the H^∞-norm of a transfer function matrix. *Syst. Control Lett.*, 14:287–293, 1990.

[5] M. J. Chen, K. N. Fan, and C. N. Nett. Structured singular values with nondiagonal structures–Part I: Characterizations, Part II: Computation. *IEEE Trans. Automat. Contr.*, 41(10):1507–1516, 1996.

[6] J. C. Doyle. Analysis of feedback systems with structured uncertainties. *IEE Proc., Part D*, 129:242–250, 1982.

[7] E. Gallestey. Computing spectral value sets using the subharmonicity of the norm of rational matrices. *BIT*, 38(1):22–33, 1998.

[8] D. Hinrichsen and B. Kelb. Spectral value sets: A graphical tool for robustness analysis. *Syst. Control Lett.*, 21:127–136, 1993.

[9] D. Hinrichsen and A. J. Pritchard. Real and complex stability radii: A survey. In D. Hinrichsen and B. Mårtensson, editors, *Control of Uncertain Systems*, pages 119–162. Birkhäuser, Basel, 1990.

[10] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness.* Springer-Verlag, Berlin, 2005.

[11] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1990.

[12] T. Hu and L. Qiu. On structured perturbation of Hermitian matrices. *Linear Algebra Appl.* 275-276, 287-314 (1998)

[13] M. Karow. *Geometry of Spectral Value Sets.* PhD thesis, University of Bremen, Germany, 2003.

[14] M. Karow. Spectral value sets for small perturbations and eigenvalue condition numbers. In preparation, Berlin University of Technology, 2007.

[15] M. Karow, D. Hinrichsen and A. J. Pritchard. Interconnected systems with uncertain couplings: explicit formulae for μ-values, spectral value sets and stability radii. *SIAM J. Control Opt.* 45(3):856-884, 2006.

[16] M. Karow and D. Kressner. The structured distance to uncontrollability. In preparation , Berlin University of Technology, 2007.

[17] M. Karow, D. Kressner, and F. Tisseur. Structured Eigenvalue Condition Numbers. *SIAM J. Matrix Anal. Appl.* 28(4):1052-1068, 2006.

[18] Craig T. Lawrence, André L. Tits, and Paul Van Dooren. A fast algorithm for the computation of an upper bound on the μ-norm. *Automatica*, 36(3):449–456, 2000.

[19] R. A. Lippert. Fixing Two Eigenvalues by a Minimal Perturbation. *Linear Algebra and its Applications* 406:177-200, 2005.

[20] N. Mackey, S. Mackey, and F. Tisseur. Structured Mapping Problems for Matrices Associated with Scalar Products Part I: Lie and Jordan Algebras. MIMS EPrint 2006.44, University of Manchester, 2006.

[21] A. N. Malyshev. A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues. Numer. Math. 83(3):443-454,1999

[22] A. Packard and J. C. Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[23] L. Qiu, B. Bernhardsson, A. Rantzer, E. J. Davison, P. M. Young, and J. C. Doyle. A formula for computation of the real stability radius. *Automatica*, 31(6):879–890, 1995.

[24] M. G. Safonov and K.-H. M. Fan. Special issue: Multivariable stability margin. *Int. J. Robust & Nonlinear Control*, 7(2):97–226, 1997.

[25] L. N. Trefethen. Computation of pseudospectra. *Acta Numerica*, 8:247–295, 1999.

[26] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators.* Princeton University Press, 2005.

[27] K. Zhou, J. C. Doyle and K. Glover. *Robust and optimal control.* Upper Saddle River, NJ: Prentice Hall, 1996.