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 quantities including stability radii, structured eigenva

1. W e now give the definition of 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 N, R, C represent the sets of positive integers, real numbers and complex numbers respectively. By C_n^m we denote the set of n by m matrices with entries in C. Furthermore, C^n = C_n^1 is the set of column vectors of length n. The conjugate transpose of A ∈ C_n^m will be written A* and A^T. If A is square then σ(A) and ρ(A) = C \ \ σ(A) denote its spectrum and its resolvent set. The n × n identity matrix will be written I_n.

By a perturbation class ∆ we mean a nonempty closed subset of C_j^\times q which is star shaped with respect to 0 ∈ C_j^\times q, i.e. if ∆ ∈ ∆ then t\Delta ∈ ∆ for 0 ≤ t ≤ 1. We now give the definition of µ-values.

DEFINITION 1.1. Let ∆ ⊆ C_j^\times q be a perturbation class and let || · || be a norm on C_j^\times q.

• The µ-value of M ∈ C_j^\times q with respect to ∆ and || · || is

\[ \mu_\Delta(M) := ( \inf \{ ||\Delta|| ; \Delta \in \Delta, 1 \in \sigma(\Delta M) \} )^{-1}. \] (1.1)

Thus \( \mu_\Delta(M) \) is the inverse of the smallest norm of a \( \Delta \in \Delta \) such that 1 is an eigenvalue of the matrix product \( \Delta M \). If there is no such \( \Delta \in \Delta \) then \( \mu_\Delta(M) = 0 \).

• If \( l = q \) then the µ-value of \( M \) of second kind is defined as

\[ \tilde{\mu}_\Delta(M) := \{ ||\Delta|| ; \Delta \in \Delta, \det(M - \Delta) = 0 \}. \] (1.2)

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

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

\[ \mu_{C_j^\times q}(M) = \sigma_{\text{max}}(M), \quad \text{and} \quad \tilde{\mu}_{C_j^\times q}(M) = \sigma_{\text{min}}(M). \] (1.3)

where \( \sigma_{\text{max}}(\cdot) \) and \( \sigma_{\text{min}}(\cdot) \) 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

\[ A \mapsto A_\Delta = A + B\Delta C, \quad \Delta \in \Delta, \quad ||\Delta|| < \delta, \] (1.4)

where \( A \in C_n^k, B \in C_n^l, C \in C^k_m \) are fixed matrices. The set of all eigenvalues of all matrices \( A_\Delta \) given by (1.4) is called a spectral value set (structured pseudospectrum). It is denoted by

\[ \sigma_{\Delta}(A, B, C; \delta) := \bigcup_{\Delta \in \Delta, ||\Delta|| < \delta} \sigma(A + B\Delta C) \]

\[ = \{ s \in \mathbb{C} ; \exists \Delta \in \Delta : ||\Delta|| < \delta, \text{ and } \det(sI_n - (A + B\Delta C)) = 0 \}. \] (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]

$$s \in \sigma(A + B\Delta C) \iff 1 \in \sigma(\Delta G(s))$$

(1.6)

it follows that

$$\mu_{\Delta}(G(s)) = \left(\inf\{ \|\Delta\| \mid \Delta \in \Delta, \ s \in \sigma(A + B\Delta C) \}\right)^{-1}, \ s \in \rho(A).$$

(1.7)

This in turn yields

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

(1.8)

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

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

(1.9)

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

$$\tilde{\mu}_{\Delta}(sI_n - A) = \inf\{ \|\Delta\| \mid \Delta \in \Delta, \ s \in \sigma(A + \Delta) \}, \ s \in \mathbb{C},$$

(1.10)

$$\sigma_{\Delta}(A; \delta) = \{s \in \mathbb{C}; \tilde{\mu}_{\Delta}(sI_n - A) < \delta\}, \ \delta > 0.$$  

(1.11)

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

The organization of this paper is as follows. In Section 2 we provide useful characterizations for $\mu$ with respect to Hermitian, symmetric and skew-symmetric perturbations. These characterizations are then used in Sections 4 and 5 to compute the associated $\mu$-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: $\mu$-values and spectral value sets for Hamiltonian perturbations of Hamiltonian matrices.

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

$$Herm(n) := \{ \Delta \in \mathbb{C}^{n \times n}; \Delta^* = \Delta \},$$

$$Sym(n) := \{ \Delta \in \mathbb{C}^{n \times n}; \Delta^T = \Delta \},$$

$$Skew(n) := \{ \Delta \in \mathbb{C}^{n \times n}; \Delta^T = -\Delta \}.$$  

(2.1)

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

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

Note that

$$\nu_{\Delta}(x, y) \geq \|y\|/\|x\| \text{ for all } x \neq 0,$$

(2.2)

since $\Delta x = y$ implies $\|\Delta\|/\|x\| \geq \|y\|$.

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

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

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

$$\tilde{\mu}_{\Delta}(M) = \inf\{ \nu_{\Delta}(v, Mv) \mid v \in \mathbb{C}^n, \ \|v\| = 1 \}.$$
Proof. This follows from the equivalences

\[ 1 \in \sigma(\Delta M) \Leftrightarrow \Delta(Mv) = v \text{ for some } v \text{ with } \|v\| = 1, \]
\[ \det(M - \Delta) = 0 \Leftrightarrow \Delta v = Mv \text{ for some } v \text{ with } \|v\| = 1. \]

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

**Proposition 2.1.** Let \( x, y \in \mathbb{C}^n \), \( x \neq 0 \). Then

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

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

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

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

(a). Let \( \alpha := x^*y \). Suppose there exists \( \Delta \in \text{Herm}(n) \) such that \( \Delta x = y \). Then \( \alpha = x^*\Delta x \in \mathbb{R} \). Thus, \( \nu_{\text{Herm}}(x, y) = \infty \) if \( \alpha \not\in \mathbb{R} \). Assume now that \( \alpha \in \mathbb{R} \) and the vectors \( x \) and \( y \) are linearly dependent. 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

\[ \Delta_0 := \alpha xx^* + \beta(zx^* + xz^*) - \alpha zz^*. \]

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

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

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

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

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

\[ \Delta_0 := yx^* - \bar{x}y^T \]

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 \( M_h = \frac{1}{2}(M - M^*) \) is definite then \( \mu_{\text{Herm}}(M) = 0 \) and \( \bar{\mu}_{\text{Herm}}(M) = \infty \).

Otherwise

\[ \mu_{\text{Herm}}(M) = \max\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1, v^*M_h v = 0 \}, \]
\[ \bar{\mu}_{\text{Herm}}(M) = \min\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1, v^*M_h v = 0 \}. \]  

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

\[ \mu_{\text{Skew}}(M) = \max\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1, v^T M_s v = 0 \} \]
\[ \bar{\mu}_{\text{Skew}}(M) = \min\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1, v^T M_s v = 0 \}. \]

(c) We always have

\[ \mu_{\text{Sym}}(M) = \max\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1 \} = \sigma_{\text{max}}(M), \]
\[ \bar{\mu}_{\text{Sym}}(M) = \min\{ \|Mv\| ; \ v \in \mathbb{C}^n, \|v\| = 1 \} = \sigma_{\text{min}}(M). \]
Proof. By (2.3) and Proposition 2.1,

\[ \mu_{\text{Herm}}(M) = (\inf \{ \nu_{\text{Herm}}(Mv, v) : v \in \mathbb{C}^n, \|v\| = 1 \})^{-1}, \]

(2.6)

\[ \nu_{\text{Herm}}(Mv, v) = \begin{cases} \|Mv\|^{-1} & \text{if } 0 = \Im((Mv)^*v) = v^*Mv \\ \infty & \text{otherwise.} \end{cases} \]

(2.7)

Hence \( \mu_{\text{Herm}}(M) = \infty^{-1} = 0 \) 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 \( v^*M_\mu v = 0 \) for some nonzero \( v \in \mathbb{C}^n \) if \( n \geq 2 \) (see Lemma 5.3). Hence \( \mu_{\text{Skew}}(M) \neq \infty \neq \tilde{\mu}_{\text{Skew}}(M). \) \( \square \)

Note that the \( \mu \)-values for the symmetric case coincide with the \( \mu \)-values for the unstructured case \( \Delta = \mathbb{C}^{n \times n} \) (see relation (1.3)). For the sake of computation of \( \mu \)-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 \( H, H_0, H_1 \in \text{Herm}(n), S \in \text{Sym}(n) \) we define

\[ \overline{m}_h(H_0, H_1) := \sup \{ v^*H_0v : v \in \mathbb{C}^n, v^*H_1v = 0, \|v\| = 1 \}, \]

(3.1)

\[ \underline{m}_h(H_0, H_1) := \inf \{ v^*H_0v : v \in \mathbb{C}^n, v^*H_1v = 0, \|v\| = 1 \}, \]

\[ \overline{m}_{hv}(H, S) := \sup \{ v^*Hv : v \in \mathbb{C}^n, v^TSv = 0, \|v\| = 1 \}, \]

\[ \underline{m}_{hv}(H, S) := \inf \{ v^*Hv : v \in \mathbb{C}^n, v^TSv = 0, \|v\| = 1 \}. \]

(2.8)

We have the following corollary to Theorem 2.2.

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

(a) if \( M_h \) is not definite,

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

\[ \tilde{\mu}_{\text{Herm}}(M) = \sqrt{\underline{m}_h(M^*M, M_h)}; \]

(b) if \( n \geq 2, \)

\[ \mu_{\text{Skew}}(M) = \sqrt{\overline{m}_{hv}(M^*M, M_s)}, \]

\[ \tilde{\mu}_{\text{Skew}}(M) = \sqrt{\underline{m}_{hv}(M^*M, M_s)}. \]

Thus, in order to calculate the \( \mu \)-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 \( H \in \text{Herm}(n) \) in decreasing order. We also use the notation \( \lambda_{\text{max}}(H) := \lambda_1(H), \lambda_{\text{min}}(H) := \lambda_n(H). \) 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 \( S_k \) denote the set of \( k \)-dimensional subspaces of \( \mathbb{C}^n. \) The Courant-Fischer principle states that

\[ \lambda_k(H) = \max_{v \in S_k} \min_{v \in \mathcal{V}} v^*Hv = \min_{v \in \mathcal{V}} \max_{v \in S_k} v^*Hv, \]

(3.1)

In particular,

\[ \lambda_{\text{max}}(H) = \max_{v \in \mathbb{C}^n} v^*Hv, \quad \lambda_{\text{min}}(H) = \min_{v \in \mathbb{C}^n} v^*Hv. \]

(3.2)

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

\[ \lambda_k(H) + \lambda_{\text{min}}(F) \leq \lambda_k(H + F) \leq \lambda_k(H) + \lambda_{\text{max}}(F), \quad H, F \in \text{Herm}(n). \]

(3.3)
Let \( \mathcal{I} \subseteq \mathbb{R} \) be an interval and let \( H : \mathcal{I} \to \text{Herm}(n) \) be an analytic function. We consider the maps

\[
\phi_k : \mathcal{I} \to \mathbb{R} \quad \phi_k(t) = \lambda_k(H(t)).
\]

Suppose that the eigenvalue \( \phi_k(\tau) \) has constant multiplicity for all \( \tau \) in a neighbourhood of \( t \in \mathcal{I} \). It is well known that in this case \( \phi_k \) 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 \( \phi_k \) 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 \( k_0 \leq j \leq k_1 \) and \( \phi_k(t) \neq \phi_j(t) \) for \( 1 \leq j < k_0 \) and \( k_1 < j \leq n \). 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

\[
\lim_{h \to 0^+} \frac{\phi_k(t+h) - \phi_k(t)}{h} = \lambda_{k_1-k_0+1}(V^* \dot{H}(t)V) = v^* \dot{H}(t)v,
\]

where \( v \in E_k(H(t)) \) is any unit vector satisfying \( v = V \xi \) for some \( \xi \in E_{k_1-k_0+1}(V^* \dot{H}(t)V) \), \( \|\xi\| = 1 \). The left derivative of \( \phi_k(\cdot) \) at \( t \) is given by

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

where \( v \in E_k(H(t)) \) is any unit vector satisfying \( v = V \xi \) for some \( \xi \in E_{k_1}(V^* \dot{H}(t)V) \), \( \|\xi\| = 1 \).

**Proof.** This follows from \([2, \text{page 149}]. \) See also \([19, \text{Theorem 8.4}]. \)

**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 \( \phi_k(t) > \phi_k(t+1) \). Then \( v^* \dot{H}(t)v \geq 0 \) 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 \([21, \text{23}]. \) 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 \( 0 \neq v \in E_k(H(t)) \) such that \( v^* \dot{H}(t)v = 0 \).

**Proof.** Assume the local extremum is a minimum. Then the right derivative of \( \phi_k \) 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 \( v_0^* \dot{H}(t)v_0 \geq 0 \) and \( v_1^* \dot{H}(t)v_1 \leq 0 \). If these inequalities are strict in both cases then \( v_0 \) and \( v_1 \) are linearly independent, and hence the vectors \( v_0 = (1- \theta ) v_0 + \theta v_1 \in E_k(H(t)) \) are nonzero for all \( \theta \in [0, 1] \). By continuity we have \( v_0^* \dot{H}(t)v_0 = 0 = v_1^* \dot{H}(t)v_1 \) for some \( \theta \). The proof for a local maximum is analogous.

### 4. Computation of \( m_k(H_0, H_1) \) and \( m_k(H_0, H_1) \). In this section we provide useful characterizations of the quantities \( m_k(H_0, H_1) \), \( m_k(H_0, H_1) \) defined in (2.8).

**Theorem 4.1.** Let \( H_0, H_1 \in \text{Herm}(n) \) 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

\[
m_k(H_0, H_1) = \sup_{t \in \mathbb{R}} \phi(t).
\]

(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

\[
t_1 = -\frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\max}(H_1)}, \quad t_2 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{|\lambda_{\min}(H_1)|}.
\]

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

\[
m_k(H_0, H_1) = \lambda_{\min}(V^* H_0 V),
\]

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 \( \{x \in \mathcal{I}; f(x) \geq c\}, c \in \mathbb{R}, \) is an interval. \( f : \mathcal{I} \to \mathbb{R} \) is said to be quasiconvex if each sublevel set \( \{x \in \mathcal{I}; f(x) \leq c\}, c \in \mathbb{R}, \) 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 \( \overline{m}_h(H_0, H_1) = \infty \).
(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).

\[
\lambda_{\min}(H_0) + \lambda_{\min}(t H_1) \leq \phi(t) \leq \lambda_{\max}(H_0) + \lambda_{\min}(t H_1),
\]
(4.4)

\[
\phi(t_*) + \lambda_{\min}(t H_1) \leq \phi(t_* + t) \leq \phi(t_*) + \lambda_{\max}(t H_1), \quad t, t_* \in \mathbb{R}.
\]
(4.5)

Note that

\[
\lambda_{\min}(t H_1) = \begin{cases} 
\lambda_{\min}(H_1) t & \text{if } t \geq 0, \\
\lambda_{\max}(H_1) t & \text{if } t \leq 0.
\end{cases}
\]

For any unit vector \( v \in \mathbb{C}^n \) satisfying \( v^* H_1 v = 0 \) and any \( t \in \mathbb{R} \) we have by (3.2),

\[
\phi(t) \leq v^* (H_0 + t H_1) v = v^* H_0 v.
\]

This implies

\[
\sup_{t \in \mathbb{R}} \phi(t) = \phi(t_0) = \overline{m}_h(H_0, H_1).
\]

(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 \( v_0^* H_1 v_0 = 0 \), whence \( v_0^* H_0 v_0 = \phi(t_0) \). Thus \( \overline{m}_h(H_0, H_1) \leq \phi(t_0) \leq \sup_{t \in \mathbb{R}} \phi(t) \), and then (4.6) yields

\[
\sup_{t \in \mathbb{R}} \phi(t) = \phi(t_0) = \overline{m}_h(H_0, H_1).
\]

Thus each local extremum of \( \phi \) is the global maximum. This implies that the function \( \phi \) is quasiconcave. Suppose now, that \( H_1 \) is indefinite and let \( t_1, t_2 \) be defined as in (4.2). Then for \( t \notin [t_1, t_2] \), \( \lambda_{\max}(H_0) + \lambda_{\min}(t H_1) \leq \lambda_{\min}(H_0) = \phi(0) \). By combining this with (4.4) we obtain \( \phi(t) \leq \phi(0) \). Consequently, \( \phi \) attains a local maximum at some \( t_0 \in [t_1, t_2] \). We thus have shown (ii) and (i) for the case that \( H_1 \) is indefinite. Suppose now, that \( H_1 \) is semidefinite. Then \( v^* H_1 v = 0 \) if and only if \( v \in \ker H_1 \). If \( H_1 \) is definite then \( \ker H_1 = \{0\} \). This implies (iii) and (iv). (vi) 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

\[
v_k^* (H_0 + k H_1) v_k = \phi(k) < \overline{m}_h(H_0, H_1) - \epsilon,
\]
whence

\[
0 \leq v_k^* H_1 v_k \leq \frac{1}{k} (\overline{m}_h(H_0, H_1) - v_k^* H_0 v_k - \epsilon).
\]

(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 \( \hat{v}^* H_0 \hat{v} \geq \overline{m}_h(H_0, H_1) \) by definition of \( \overline{m}_h(H_0, H_1) \). Therefore \( \overline{m}_h(H_0, H_1) - v_{k_j}^* H_0 v_{k_j} - \epsilon < 0 \) for \( k_j \) large enough. The latter contradicts (4.7).

Remark 4.2. The function

\[
(H_0, H_1) \mapsto \overline{m}_h(H_0, H_1)
\]

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 \( E_1 \in \mathcal{U}_c = \{ E \in \text{Herm}(n); \|E\| \leq \epsilon, \epsilon > 0 \} \). Let

\[
a = \max \{ \lambda_{\max}(H_0 + E) - \lambda_{\min}(H_0 + E); \ E \in \mathcal{U}_c \},
\]

\[
T_1 = -a/\min \{ \lambda_{\max}(H_1 + E); \ E \in \mathcal{U}_c \},
\]

\[
T_2 = a/\min \{ |\lambda_{\min}(H_1 + E)|; \ E \in \mathcal{U}_c \}.
\]
Then by part (ii) of Theorem 4.1 we have for all \((E_0, E_1) \in \mathcal{U}_t^2\),
\[
\mu_2(H_0 + E_0, H_1 + E_1) = \max_{t \in [t_1, t_2]} \lambda_{\min}(H_0 + E_0 + t(H_1 + E_1)).
\]

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 \(H_0, H_1 \in \text{Herm}(n)\). 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 \(\mu_2(H_0, H_1)\) 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 \(\mu_2(H_0, H_1)\) we have the following result. Its proof is analogous to that of Theorem 4.1. and therefore omitted.

Theorem 4.5. Let \(H_0, H_1 \in \text{Herm}(n)\) and \(\phi(t) = \lambda_{\max}(H_0 + tH_1), t \in \mathbb{R}\).

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

\[
\mu_2(H_0, H_1) = \inf_{t \in \mathbb{R}} \phi(t). \tag{4.9}
\]

(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

\[
t_1 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\min}(H_1)}, \quad t_2 = \frac{\lambda_{\max}(H_0) - \lambda_{\min}(H_0)}{\lambda_{\max}(H_1)}. \tag{4.10}
\]

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

\[
\mu_2(H_0, H_1) = \lambda_{\max}(V^*H_0V), \tag{4.11}
\]

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

(iv) If \(H_1\) is definite then \(\mu_2(H_0, H_1) = -\infty\).

(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 \(\lim_{t \to -\infty} \phi(t) = -\infty\) (\(\lim_{t \to -\infty} \phi(t) = -\infty\)).

5. Computation of \(\mu_2(H, S)\) and \(\mu_2(H, S)\). We now treat the Hermitian-symmetric case. Recall the definition in (2.8),

\[
\mu_2(H, S) = \sup \{ v^*Hv ; v \in \mathbb{C}^n, v^*Sv = 0, \|v\| = 1 \}, \quad H \in \text{Herm}(n), S \in \text{Sym}(n).
\]

If \(S = 0\) then \(\mu_2(H, S) = \lambda_{\max}(H)\). Suppose that \(\text{rank}(S) = 1\). Then \(S\) can be written in the form \(S = xx^\top\) for some nonzero \(x \in \mathbb{C}^n\). Hence, for any \(v \in \mathbb{C}^n, v^*Sv = (v^*x)^2\). Let \(\mathcal{V} = \{ v \in \mathbb{C}^n ; v^*x = 0 \}\). Then

\[
\mu_2(H, S) = \mu_2(H, xx^\top) = \sup_{v \in \mathcal{V}, \|v\| = 1} v^*Hv. \tag{5.1}
\]

If \(n = 1\) then \(\mathcal{V} = 0\). Thus \(\mu_2(H, xx^\top) = -\infty\). 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

\[
\mu_2(H, xx^\top) = \max_{\xi \in \mathbb{C}^{n-1}, \|\xi\| = 1} (V\xi)^*H(V\xi) = \lambda_{\max}(V^*HV).
\]

If \(\text{rank}(S) \geq 2\) then \(\mu_2(H, S)\) can be computed by minimizing the second largest eigenvalue of a hermitian pencil. The precise statement is as follows.

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

\[
\mu_2(H, S) = \min_{0 \leq t \leq t_1} \lambda_2 \left( \begin{bmatrix} H & tS \\ t^\top & t^\top \end{bmatrix} \right).
\]

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
\[
F(t) := \begin{bmatrix} H & tS \\ tS & t \overline{H} \end{bmatrix}, \quad \phi(t) := \lambda_2(F(t)).
\]

• For a unit vector \( v \in \mathbb{C}^n \),
\[
\mathcal{U}_v := \left\{ \begin{bmatrix} z_1v \\ z_2v \end{bmatrix} : z_1, z_2 \in \mathbb{C} \right\}.
\]
Note that \( \mathcal{U}_v \) is a 2-dimensional subspace of \( \mathbb{C}^{2n} \), and
\[
F(t) \begin{bmatrix} z_1v \\ z_2v \end{bmatrix} = (|z_1|^2 + |z_2|^2) v^* Hv + 2t \Re(z_1z_2 v^\top S v), \quad z_1, z_2 \in \mathbb{C}. \tag{5.2}
\]

**Lemma 5.1.** For any \( H \in \text{Her}_n \), \( S \in \text{Sym}_n \) we have \( \overline{m}_{hs}(H, S) \leq \inf_{t \in \mathbb{R}} \phi(t) \).

*Proof.* Let \( v \in \mathbb{C}^n \) be a unit vector satisfying \( v^\top S v = 0 \). Then by the Courant-Fischer max-min-principle and (5.2),
\[
\phi(t) \geq \min_{x \in \mathcal{U}_v, \|x\|=1} x^* H x = v^* H v \quad \text{for all } t \in \mathbb{R}.
\]
Hence, \( \phi(t) \geq \overline{m}_{hs}(H, S) \).

**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 \( \dim \mathcal{V} \geq 2 \). Then to any \( S \in \text{Sym}_n \) 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 \( (z_1, z_2) \neq (0,0) \).

**Lemma 5.4.** The following statements are equivalent.

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

(ii) Either \( \dim E_1(H) \geq 2 \), 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 \( \overline{m}_{hs}(H, S) \leq \max \{ v^* H v : v \in \mathbb{C}^n, \|v\|=1 \} = \lambda_{\max}(H) \). 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 \( \dim E_1(H) \geq 2 \).

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

(iii) \( \Rightarrow \) (i). Since (i) is satisfied if \( \dim E_1(H) \geq 2 \) 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 \( 0 \leq x^* F(0) x \) for all \( x \in E_2(F(0)) \). In other words, we have for all \( z_1, z_2 \in \mathbb{C} \),
\[
0 \leq \begin{bmatrix} z_1v \\ z_2v \end{bmatrix}^* \begin{bmatrix} 0 & \overline{S} \\ S & 0 \end{bmatrix} \begin{bmatrix} z_1v \\ z_2v \end{bmatrix} = 2 \Re(z_1z_2 v^\top S v).
\]
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 \( t_0 \neq 0 \). 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, \( \phi(t_0) \leq \overline{m}_{hs}(H, S) \).

*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
\[
F(t_0) v_0 = \phi(t_0) v_0, \quad v_0^* \overline{F}(t_0) v_0 = 0. \tag{5.3}
\]

\[
F(t_0) v_0 = \phi(t_0) v_0, \quad v_0^* \overline{F}(t_0) v_0 = 0. \tag{5.4}
\]
Let
\[ H_0 := H - \phi(t_0) I_n, \quad v_0 = \begin{bmatrix} x \\ y \end{bmatrix}, \quad x, y \in \mathbb{C}^n. \]

Then (5.3) is equivalent to the equations
\[ H_0 x = -t_0 S y, \quad H_0 y = -t_0 S x, \]
which imply
\[ x^* H_0 x = -t_0 x^* S y = y^* H_0 y, \quad x^* H_0 y = -t_0 x^* S x = -t_0 y^* S y. \]

Since \( t_0 \neq 0 \) it follows that
\[ x^* S y \in \mathbb{R}, \quad y^* S x = x^* S x. \]

We have that \( v_0^* F(t_0)v_0 = 2 \Re(x^* S y) \). Thus, (5.4) and (5.7) yield
\[ x^* S y = 0. \tag{5.9} \]

Now let
\[ \beta := \begin{cases} 1 & \text{if } x^* S x = 0, \\ \frac{2}{x^* S x} & \text{otherwise}. \end{cases} \]

Then we have
\[
(x \pm \beta y)^* S (x \pm \beta y) = x^* S x + \beta^2 y^* S y + 2 \beta x^* S y \\
= x^* S x + \beta^2 \overline{x^* S x} + 2 \beta \overline{x^* S y} \\
= -x^* S x = 0
\]
and
\[
(x \pm \beta y)^* H_0 (x \pm \beta y) = x^* H_0 x + |\beta|^2 y^* H_0 y + 2 \Re(x^* H_0 y \beta) \\
= -t_0 \left( 1 + |\beta|^2 \right) x^* S y + 2 \Re(x^* S x \beta) = 0
\]
(using (5.6) and (5.9))

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 \( \text{rank}(S) \geq 2 \) then \( \phi \) attains its minimum in the interval \([0, t_1]\), where \( t_1 = \frac{2\|H\|}{\sigma_2(S)} \).

**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 \( \begin{bmatrix} 0 & t S \\ t S & 0 \end{bmatrix} \) are the singular values of \( S \) and their negatives. In particular,
\[ \lambda_2 \left( \begin{bmatrix} 0 & t S \\ t S & 0 \end{bmatrix} \right) = \sigma_2(tS) = |t| \sigma_2(S). \]

Furthermore,
\[ \phi(0) = \lambda_2 \left( \begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix} \right) = \lambda_{\text{max}} \left( \begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix} \right) = \lambda_{\text{max}}(H) \leq \|H\|. \]
We conclude that
\[ \phi(t) = \lambda_2 \left( \begin{bmatrix} H & tS \\ t^*S & \Pi \end{bmatrix} \right) \geq \lambda_2 \left( \begin{bmatrix} 0 & tS \\ t^*S & 0 \end{bmatrix} \right) - \lambda_{\max} \left( \begin{bmatrix} H & 0 \\ 0 & \Pi \end{bmatrix} \right) \geq |t| \sigma_2(S) - \|H\|. \]

Thus, if \(|t| \geq t_1\) then \(\phi(t) \geq \phi(0)\). Consequently \(\phi\) 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 \(t_0 \geq 0\). \(\square\)

We now summarize the results and finish the proof of Theorem 5.1. By Lemma 5.1,
\[ m_{hs}(H, S) \leq \inf_{t \in \mathbb{R}} \phi(t). \quad (5.10) \]

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 \(\phi\) is quasiconvex, and the proof is complete.

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

**Theorem 5.2.** Let \(H \in \text{Herm}(n)\), \(S \in \text{Sym}(n)\) with \(\text{rank}(S) \geq 2\), and \(t_1 = 2\frac{\|H\|}{\sigma_2(S)}\). Then
\[ m_{hs}(H, S) = \max_{0 \leq t \leq t_1} \lambda_{2n-1} \left( \begin{bmatrix} H & tS \\ t^*S & \Pi \end{bmatrix} \right). \]

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
\[
\begin{align*}
\text{Herm}(n) \times \text{Sym}(n) \ni (H, S) &\mapsto m_{hs}(H, S), \\
\text{Herm}(n) \times \text{Sym}(n) \ni (H, S) &\mapsto m_{hs}(H, S)
\end{align*}
\]
are continuous at all \((H, S)\) with \(\text{rank}(S) \geq 2\).

**6. Self- and skew-adjoint matrices.** We now treat \(\mu\)-values with respect to linear subspaces which are induced by a scalar product on \(\mathbb{C}^n\). Specifically we show that these \(\mu\)-values are closely related to the \(\mu\)-values with respect to Hermitian, symmetric and skew-symmetric perturbations.

For nonsingular \(\Pi \in \mathbb{C}^{n \times n}\) we consider the scalar products
\[ \langle x, y \rangle_{\Pi} = x^* \Pi y, \quad x, y \in \mathbb{C}^n, \quad * \in \{*, \top\}. \]

Depending on whether \(* = \top\) or \(* = *\) the scalar product is a bilinear form or a sesquilinear form. We assume that \(\Pi\) satisfies a symmetry relation of the form
\[ \Pi^* = \epsilon_0 \Pi, \quad \text{with} \quad \epsilon_0 = -1 \text{ or } \epsilon_0 = 1. \quad (6.1) \]

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
\[ \langle \Delta x, y \rangle_{\Pi} = \epsilon \langle x, \Delta y \rangle_{\Pi} \quad \text{for all } x, y \in \mathbb{C}^n, \quad \epsilon = 1 (\epsilon = -1). \quad (6.2) \]

and \(\epsilon = 1 (\epsilon = -1)\). The relation (6.2) is easily seen to be equivalent to
\[ \Delta^* \Pi = \epsilon \Pi \Delta. \quad (6.3) \]

We denote the sets of self- and skew-adjoint matrices by
\[ \text{struct}(\Pi, *, \epsilon) := \{ \Delta \in \mathbb{C}^{n \times n} ; \Delta^* \Pi = \epsilon \Pi \Delta \}. \]

The relation (6.1) implies that (6.3) is equivalent to
\[ (\Pi \Delta)^* = \epsilon_0 \epsilon \Pi \Delta. \quad (6.4) \]

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

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

In many applications $\Pi$ is unitary. The most common examples are $\Pi \in \{\text{diag}(I_n, -I_n), E_n, J_n\}$, where

$$J_n := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{C}^{2n \times 2n}, \quad E_n := \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{C}^{n \times n}. \quad (6.5)$$

For unitary $\Pi$ the $\mu$-values of the associated self- and skew-adjoint classes can be expressed in terms of the $\mu$-values for $\text{Herm}(n)$, $\text{Sym}(n)$ and $\text{Skew}(n)$:

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

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

and

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

Proof. Since $\Pi$ is unitary, we have

$$\mu_{\text{struct}}(M) = (\inf \{ \| \Delta \| : \Delta \in \text{struct}, \quad 1 \in \sigma(\Delta M) \})^{-1} = (\inf \{ \| \Pi \Delta \| : \Delta \in \text{struct}, \quad 1 \in \sigma((\Pi \Delta)(M \Pi^*)) \})^{-1}, \quad (6.8)$$

$$\tilde{\mu}_{\text{struct}}(M) = \inf \{ \| \Delta \| : \Delta \in \text{struct}, \quad \det(M - \Delta) = 0 \} = \inf \{ \| \Pi \Delta \| : \Delta \in \text{struct}, \quad \det(\Pi M - \Pi \Delta) = 0 \}. \quad (6.9)$$

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) $\Pi$ by $\pm i \Pi$ one obtains the fourth identity in (6.6) and (6.7) from the fourth equivalence in Lemma 6.1. \( \square \)

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

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

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

$$H = \begin{bmatrix} A & C \\ B & -A^* \end{bmatrix} \quad \text{with } A \in \mathbb{C}^{n \times n} \text{ and } B, C \in \text{Herm}(n).$$
The spectral value sets of \( H \) with respect to unstructured perturbations are by (1.3) and (1.11),
\[
\sigma_{C_n \times n}(H; \delta) = \bigcup_{\Delta \in C_n \times n, \|\Delta\| < \delta} \sigma(H + \Delta) = \{ s \in C; \sigma_{\min}(s I_n - H) < \delta \}, \quad \delta > 0.
\]

The spectral value sets with respect to Hamiltonian perturbations are
\[
\sigma_{\text{Ham}}(H; \delta) = \bigcup_{\Delta \in \text{Ham}(n), \|\Delta\| < \delta} \sigma(H + \Delta) = \{ s \in C; f(s) < \delta \}, \quad \delta > 0,
\]
where
\[
f : C \to \mathbb{R}, \quad f(s) := \tilde{\mu}_{\text{Ham}}(s I_n - H).
\]

By Corollary 6.1 we have \( f(s) = \tilde{\mu}_{\text{Herm}}(\Phi(s)) \), where \( \Phi(s) = J_n(s I_n - H) = \begin{bmatrix} -B & s I_n + A^* \\ -s I_n + A & C \end{bmatrix} \).

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,
\[
f(s) = \sqrt{m_h(\Phi(s)^* \Phi(s), \frac{\Phi(s) - \Phi(s)}{2i})} \\
= \sqrt{m_h((s I_n - H)^*(s I_n - H), -i(\Re s) J_n)} \\
= \begin{cases} \\
\sqrt{m_h((s I_n - H)^*(s I_n - H), 0)} & \text{if } s \in \mathbb{R}, \\
\sqrt{m_h((s I_n - H)^*(s I_n - H), i J_n)} & \text{otherwise,}
\end{cases}
\]
\[
= \begin{cases} \\
\sigma_{\min}(s I_n - H) & \text{if } s \in \mathbb{R}, \\
\sqrt{\max_{t \in \mathbb{R}} \lambda_{\min}((s I_n - H)^*(s I_n - H) + t i J_n)} & \text{otherwise.}
\end{cases}
\]

\[\text{Fig. 7.1. The sets } \sigma_{\text{Ham}}(H; 1) \text{ (upper row) and } \sigma_{C_n \times n}(H; 1) \text{ (lower row).}\]
Furthermore, by (4.2) the maximum in (7.3) is attained for some \( t \in \mathbb{R} \) satisfying
\[
|t| \leq \sigma_{\max}(sI_n - H)^2 - \sigma_{\min}(sI_n - H)^2.
\]

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

**Proposition 7.2.** Let \( H \in \mathcal{H}_{s I}(n) \), 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 \( \sigma_{\Ham}(H; \delta) \) is an open subset of \( \mathbb{C} \).

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

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

(a) The function \( f \) is discontinuous at \( s \).

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

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

**Proof.** From Remark 4.2 we obtain that the function
\[
g : \mathbb{C} \to \mathbb{R}, \quad g(s) = \sqrt{m_6((sI_n - H)^* (sI_n - H), iJ_n)}
\]
is continuous. By definition of \( m_6 \) we have for all \( s \in \mathbb{C} \) that \( g(s) \leq \sigma_{\min}(sI_n - H) \). Equality holds iff there exists a right singular vector \( v \neq 0 \) to the singular value \( \sigma_{\min}(sI_n - H) \) such that \( v^*J_nv = 0 \). If \( s \) is an eigenvalue of \( H \) then the singular vectors to \( \sigma_{\min}(sI_n - H) = 0 \) 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 \( f(s) = \sigma_{\min}(sI_n - H) \) 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
\[
H_\gamma = \begin{bmatrix} 0 & \gamma^{-1}C \\ \gamma B & 0 \end{bmatrix}, \quad B = \text{diag}(1, 6, -6), \quad C = \text{diag}(1, 6, -6), \quad \gamma \in \{1, 1.3, 5, 6\}.
\]
The lower row in the figure shows the sets \( \sigma_{\mathcal{C}^{\infty} s I}(H; \gamma; 1) \) for comparison. The crosses mark the eigenvalues of \( H_\gamma \).

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

**REFERENCES**


