ON SYLVESTER'S LAW OF INERTIA FOR NONLINEAR EIGENVALUE PROBLEMS

ALEKSANDRA KOSTIĆ* AND HEINRICH VOSS†

Key words. eigenvalue, variational characterization, minmax principle, Sylvester's law of inertia

AMS subject classification. 15A18, 65F15

Abstract. For Hermitian matrices and generalized definite eigenproblems the LDL^H factorization yields an easy tool to slice the spectrum into two disjunct intervals. In this note we generalize this method to nonlinear eigenvalue problems allowing for a minmax characterization of (some of) its real eigenvalues. In particular we apply this approach to several classes of quadratic pencils.

1. Introduction.

The inertia of a Hermitian matrix A is the triplet of nonnegative integers $\mathrm{In}(A) := (n_p, n_n, n_z)$ where n_p, n_n, and n_z are the number of positive, negative, and zero eigenvalues of A (counting multiplicities). Sylvester's classical law of inertia states that two Hermitian matrices $A, B \in \mathbb{C}^{n\times n}$ are congruent (i.e. $A = S^H BS$ for some nonsingular matrix S) if and only if they have the same inertia $\mathrm{In}(A) = \mathrm{In}(B)$.

An obvious consequence of the law of inertia is the following corollary: If $A = LDL^H$ is the LDL^H factorization of A, then n_p and n_n equals the number of positive and negative entries of D, respectively. Hence, the inertia of a matrix can be computed easily (in particular if the matrix is banded).

If $B \in \mathbb{C}^{n\times n}$ is positive definite, $A - \sigma B = LDL^H$ for some $\sigma \in \mathbb{R}$, and D has n_n negative elements, then the generalized eigenvalue problem $Ax = \lambda Bx$ has n_n eigenvalues smaller than σ. Hence, the law of inertia yields a tool to locate eigenvalues of Hermitian matrices or definite matrix pencils. Combining it with bisection or the secant methods one can determine all eigenvalues in a given interval or determine initial approximations for fast eigensolvers, and it can be used to check whether a method has found all eigenvalues in an interval of interest or not.

The law of inertia was first proved 1858 by J.J. Sylvester [19], and several different proofs can be found in textbooks [3, 6, 11, 13, 15], one of which is based on the minmax characterization of eigenvalues of Hermitian matrices. In this note we discuss generalizations of the law of inertia to nonlinear eigenvalue problems allowing for a minmax characterization of its eigenvalues.

2. Minmax characterization.

Our main tools in this paper are variational characterizations of eigenvalues of nonlinear eigenvalue problems generalizing the well known minmax characterization of Poincaré [16] or Courant [2] and Fischer [5] for linear eigenvalue problems.

We consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0, \tag{2.1}$$

where $T(\lambda) \in \mathbb{C}^{n\times n}$, $\lambda \in J$, is a family of Hermitian matrices depending continuously on the parameter $\lambda \in J$, and J is a real open interval which may be unbounded.

---

*Mašinski fakultet Sarajevo, Univerzitet Sarajevo, Bosnia and Herzegovina, (`kostic@mef.unsa.ba`)

†Institute of Mathematics, Hamburg University of Technology, D-21071 Hamburg, Germany (`voss@tu-harburg.de`)

To generalize the variational characterization of eigenvalues we need a generalization of the Rayleigh quotient. To this end we assume that

(A_1) for every fixed $x \in \mathbb{C}^n$, $x \neq 0$ the real equation

$$f(\lambda; x) := x^H T(\lambda) x = 0 \tag{2.2}$$

has at most one solution $p(x) \in J$.

Then $f(\lambda; x) = 0$ implicitly defines a functional p on some subset $\mathcal{D} \subset \mathbb{C}^n$ which is called the Rayleigh functional of (2.1), and which is exactly the Rayleigh quotient in case of a monic linear matrix function $T(\lambda) = \lambda I - A$.

Generalizing the definiteness requirement for linear pencils $T(\lambda) = \lambda B - A$ we further assume that

(A_2) for every $x \in \mathcal{D}$ and every $\lambda \in J$ with $\lambda \neq p(x)$ it holds that

$$(\lambda - p(x)) f(\lambda; x) > 0. \tag{2.3}$$

If p is defined on $\mathcal{D} = \mathbb{C}^n \setminus \{0\}$ then the problem $T(\lambda)x = 0$ is called overdamped. This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda Cx + Kx = 0 \tag{2.4}$$

where M, C and K are Hermitian and positive definite matrices. If C is big enough such that $d(x) := (x^H Cx)^2 - 4(x^H Kx)(x^H Mx) > 0$ for every $x \neq 0$ then $T(\cdot)$ is overdamped. Generalizations of the minmax and maxmin characterizations of eigenvalues where proved by Duffin [4] for the quadratic case and by Rogers [17] for general overdamped problems.

For nonoverdamped eigenproblems the natural ordering to call the smallest eigenvalue the first one, the second smallest the second one, etc., is not appropriate. This is obvious if we make a linear eigenvalue $T(\lambda)x := (\lambda I - A)x = 0$ nonlinear by restricting it to an interval J which does not contain the smallest eigenvalue of A. Then the conditions (A_1) and (A_2) are satisfied, p is the restriction of the Rayleigh quotient R_A to

$$\mathcal{D} := \{x \neq 0 \,:\, R_A(x) \in J\},$$

and $\inf_{x \in \mathcal{D}} p(x)$ will in general not be an eigenvalue.

If $\lambda \in J$ is an eigenvalue of $T(\cdot)$ then $\mu = 0$ is an eigenvalue of the linear problem $T(\lambda)y = \mu y$, and therefore there exists $\ell \in \mathbb{N}$ such that

$$0 = \max_{V \in H_\ell} \min_{v \in V \setminus \{0\}} \frac{v^H T(\lambda) v}{\|v\|^2}$$

where H_ℓ denotes the set of all ℓ–dimensional subspaces of $\mathbb{C}^n$. In this case λ is called an ℓ-th eigenvalue of $T(\cdot)$.

With this enumeration the following minmax characterization for eigenvalues was proved in [20, 21]

THEOREM 2.1. *Let J be an open interval in $\mathbb{R}$, and let $T(\lambda) \in \mathbb{C}^{n \times n}$, $\lambda \in J$, be a family of Hermitian matrices depending continuously on the parameter $\lambda \in J$ such that the conditions (A_1) and (A_2) are satisfied.*

(i) For every $\ell \in \mathbb{N}$ there is at most one ℓ-th eigenvalue of $T(\cdot)$ which can be characterized by

$$\lambda_\ell = \min_{V \in H_\ell,\ V \cap \mathcal{D} \neq \emptyset} \ \sup_{v \in V \cap \mathcal{D}} p(v). \tag{2.5}$$

(ii) If

$$\lambda_\ell := \inf_{V\in H_\ell,\ V\cap\mathcal{D}\neq\emptyset}\ \sup_{v\in V\cap\mathcal{D}} p(v) \in J$$

for some $\ell \in \mathbb{N}$ *then* λ_ℓ *is the* ℓ*-th eigenvalue of* $T(\cdot)$ *in* J*, and* (2.5) *holds.*

(iii) Let $\lambda_1 = \inf_{x\in\mathcal{D}} p(x) \in J$ *and* $\lambda_\ell \in J$*. If the minimum in* (2.5) *is attained for an* ℓ *dimensional subspace* V*, then* $V \subset \mathcal{D}\cup\{0\}$*, and* (2.5) *can be replaced with*

$$\lambda_\ell = \min_{V\in H_\ell,\ V\subset\mathcal{D}\cup\{0\}}\ \sup_{v\in V,\ v\neq 0} p(v). \tag{2.6}$$

(iv) $\tilde\lambda$ *is an* ℓ*-th eigenvalue if and only if* $\mu = 0$ *is the* ℓ *largest eigenvalue of the linear eigenproblem* $T(\tilde\lambda)x = \mu x$*.*

(v) The minimum in (2.5) *is attained for the invariant subspace of* $T(\lambda_\ell)$ *corresponding to its* ℓ *largest eigenvalues.*

3. Sylvester's law for nonlinear eigenvalue problems.

We first consider the overdamped case.

THEOREM 3.1. *Assume that* $T : J \to \mathbb{C}^{n\times n}$ *satisfies the conditions of the minmax characterization in Theorem 2.1, and assume that the nonlinear eigenvalue problem* (2.1) *is overdamped, i.e. for every* $x \neq 0$ *equation* (2.2) *has a unique solution* $p(x) \in J$*.*

For $\sigma \in J$ *let* (n_p, n_n, n_z) *be the inertia of* $T(\sigma)$*. Then the nonlinear eigenproblem* $T(\lambda)x = 0$ *has* n *eigenvalues in* J*,* n_p *of which are smaller than* σ*,* n_n *exceed* σ*, and for* $n_z > 0$*,* σ *is an eigenvalue of multiplicity* n_z*.*

Proof. The invariant subspace W of $T(\sigma)$ corresponding to its positive eigenvalues has dimension n_p, and it holds that $f(\sigma; x) = x^H T(\sigma)x > 0$ for every $x \in W$, $x \neq 0$. Hence, $p(x) < \sigma$ by (A_2), and therefore the n_p smallest eigenvalue of $T(\cdot)$ satisfies

$$\lambda_{n_p} = \min_{\dim V = n_p}\ \max_{x\in V, x\neq 0} p(x) \le \max_{x\in W, x\neq 0} p(x) < \sigma.$$

On the other hand for every subspace V of $\mathbb{C}^n$ of dimension $n_p + n_z + 1$ there exists $x \in V$ such that $f(\sigma; x) < 0$. Thus $p(x) > \sigma$, and

$$\lambda_{n_p+n_z+1} = \min_{\dim V = n_p+n_z+1}\ \max_{x\in V, x\neq 0} p(x) > \sigma,$$

which completes the proof. □

Next we consider the case that an extreme eigenvalue $\lambda_1 := \inf_{x\in\mathcal{D}} p(x)$ or $\lambda_n := \sup_{x\in\mathcal{D}} p(x)$ is contained in J.

THEOREM 3.2. *Assume that* $T : J \to \mathbb{C}^{n\times n}$ *satisfies the conditions of the minmax characterization, and let* (n_p, n_n, n_z) *be the inertia of* $T(\sigma)$ *for some* $\sigma \in J$*.*

(i) If $\lambda_1 := \inf_{x\in\mathcal{D}} p(x) \in J$*, then the nonlinear eigenproblem* $T(\lambda)x = 0$ *has exactly* n_p *eigenvalues* $\lambda_1 \le \cdots \le \lambda_{n_p}$ *in* J *which are less than* σ*.*

(ii) If $\sup_{x\in\mathcal{D}} p(x) \in J$*, then the nonlinear eigenproblem* $T(\lambda)x = 0$ *has exactly* n_n *eigenvalues* $\lambda_{n-n_n+1} \le \cdots \le \lambda_n$ *in* J *exceeding* σ*.*

Proof. (i): We first show that $f(\lambda; x) < 0$ for every $\lambda \in J$ with $\lambda < \lambda_1$ and for every $x \neq 0$. Assume that $f(\lambda; x) \ge 0$ for some $\lambda < \lambda_1$ and $x \neq 0$, let $\hat x$ be an eigenvector of (2.1) corresponding to λ_1, and let $w(t) := t\hat x + (1-t)x$, $0 \le t \le 1$. Then, $\phi(t) := f(\lambda; w(t))$ is continuous in $[0, 1]$, $\phi(0) = f(\lambda; x) \ge 0$, and $\phi(1) = f(\lambda; \hat x) < 0$.

Hence, there exists $\hat{t} \in [0,1)$ such that $f(\lambda; w(\hat{t})) = 0$, i.e. $w(\hat{t}) \in \mathcal{D}$ and $p(w(\hat{t})) = \lambda < \lambda_1$ contradicting $\lambda_1 := \inf_{x\in\mathcal{D}} p(x)$.

For $n_p = 0$ the matrix $T(\sigma)$ is negative semidefinit, i.e. $x^H T(\sigma)x \le 0$ for $x \ne 0$, and it follows from (A_2) that $p(x) \ge \sigma$ for every $x \in \mathcal{D}$. Hence, there is no eigenvalue less than σ.

For $n_p > 0$ let W denote the invariant subspace of $T(\sigma)$ corresponding to its positive eigenvalues. Then $f(\sigma; x) = x^H T(\sigma)x > 0$ for $x \in W$, $x \ne 0$, and from $f(\lambda; x) < 0$ for $\lambda < \lambda_1$ it follows that $x \in \mathcal{D}$ and $p(x) < \sigma$. Hence, $W \subset \mathcal{D} \cup \{0\}$ and as in the proof of Theorem 3.1 we get

$$\lambda_{n_p} = \inf_{\dim V = n_p,\ V\cap\mathcal{D}\ne\emptyset} \ \max_{x\in V\cap\mathcal{D}} \ p(x) \le \max_{x\in W, x\ne 0} p(x) < \sigma,$$

i.e. $T(\cdot)$ has at least n_p eigenvalues less than σ.

Assume that there exists a (n_p+1)th eigenvalue $\lambda_{n_p+1} < \sigma$, and let W be the invariant subspace of $T(\lambda_{n_p+1})$ corresponding to its nonnegative eigenvalues. Then $W \setminus \{0\} \subset \mathcal{D}$, and $p(x) \le \lambda_{n_p+1} < \sigma$ for every $x \in W$, $x \ne 0$, contradicting that for every subspace V with $\dim V = n_p + 1$ there exists $x \in V$ with $x^H T(\sigma)x \le 0$, i.e. $x \notin \mathcal{D}$ or $p(x) \ge \sigma$.

(ii) $S(\lambda) := -T(-\lambda)$ satisfies the conditions of Theorem 2.1 in in the interval $-J$, and $-J$ contains the smallest eigenvalue $-\lambda_n$ of S. □

For the general case the law of inertia obtains the following form:

THEOREM 3.3. *Let $T : J \to \mathbb{C}^{n\times n}$ satisfy the conditions of the minmax characterization, and let $\sigma, \tau \in J$, $\sigma < \tau$.*

Let $(n_{p_\sigma}, n_{n_\sigma}, n_{z_\sigma})$ and $(n_{p_\tau}, n_{n_\tau}, n_{z_\tau})$ the inertia of $T(\sigma)$ and $T(\tau)$, respectively. Then the eigenvalue problem (2.1) has exactly $n_{p_\tau} - n_{p_\sigma}$ eigenvalues $\lambda_{n_{p_\sigma}+1} \le \dots \le \lambda_{n_{p_\tau}}$ in (σ, τ).

Proof. Let V be a subspace of $\mathbb{C}^n$ with $V \cap \mathcal{D} \ne \emptyset$ and $\dim V = n_{p_\sigma} + 1$. We first show that there exits $x \in V \cap \mathcal{D}$ with $p(x) > \sigma$ from which we then obtain

$$\lambda_{n_{p_\sigma}+1} := \inf_{dim V = n_{p_\sigma}+1,\ V\cap\mathcal{D}\ne\emptyset} \ \sup_{x\in V\cap\mathcal{D}} \ p(x) > \sigma.$$

From $\dim V > n_{p_\sigma}$ it follow that there exists $x \in V$, $x \ne 0$ such that $x^H T(\sigma)x < 0$. If $x \in \mathcal{D}$ then it follows from (A_2) that we are done. Otherwise, we choose $y \in V \cap \mathcal{D}$ and $\omega > \min(p(y), \sigma)$. Then $x^H T(\omega)x < 0 < y^H T(\omega)y$, and with $w(t) := tx + (1-t)y$ it follows in the same way as in the proof of Theorem 3.2 that there exist $\hat{t} \in [0,1]$ such that $w(\hat{t}) \in \mathcal{D}$ and $p(w(\hat{t})) = \omega > \sigma$.

If U denotes the invariant subspace of $T(\tau)$ corresponding to its positive eigenvalues, then $x^H T(\tau)x > 0$, and therefore $p(x) < \tau$ for every $x \in U \cap \mathcal{D}$. Form $\tau \ge \sigma$ we get $U \cap \mathcal{D} \ne \emptyset$, and

$$\lambda_{n_{p_\tau}} = \inf_{\dim V = n_{p_\tau},\ V\cap\mathcal{D}\ne\emptyset} \ \sup_{x\in V\cap\mathcal{D}} \ p(x) \le \sup_{x\in U\cap\mathcal{D}} \ p(x) < \tau.$$

Hence, $\lambda_{n_{p_\sigma}+1}$ and $\lambda_{n_{p_\tau}}$ are both contained in (σ, τ), and so are the eigenvalues λ_j, $j = n_{p_\sigma} + 1, \dots, n_{p_\tau}$. □

REMARK 3.4. *Not using the minmax characterization of eigenvalues Neumaier [13] proved Theorem 3.3 for matrices $T : J \to \mathbb{C}^{n\times n}$ which are Hermitian and (elementwise) differentiable in J with positive definite derivative $T'(\lambda)$, $\lambda \in J$. Obviously, such $T(\cdot)$ satisfies the conditions of the minmax characterization.*

EXAMPLE 3.5. *Consider the rational eigenvalue problem*

$$T(\lambda) := -K + \lambda M + \sum_{j=1}^{p} \frac{\lambda}{\sigma_j - \lambda} C_j C_j^T, \tag{3.1}$$

where $K, M \in \mathbb{R}^{n\times n}$ *are symmetric and positive definite,* $C_j \in \mathbb{R}^{n\times k_j}$ *has rank* k_j*, and* $\sigma_1 < \cdots < \sigma_p$*, which models the free vibrations of certain fluid–solid structures (cf. [1]).*

In each interval $J_\ell := (\sigma_\ell, \sigma_{\ell+1})$*,* $\ell = 0, \dots, p+1$*,* $\sigma_0 = 0$*,* $\sigma_{p+1} = \infty$ *the function* $f_\ell(\lambda, x) := x^H T(\lambda) x$ *is strictly monotonically increasing, and therefore all eigenvalues in* J_ℓ *are minmax values of the Rayleigh functional* p_ℓ*.*

For the first interval J_0 *Theorem 3.2 applies. Hence, if* $\tau \in J_0$ *and* (n_p, n_n, n_z) *is the inertia of* $T(\tau)$*, then there are exactly* n_p *eigenvalues in* J_0 *which are less than* τ*. Moreover, if* $\tau_1 < \tau_2$ *are contained in one interval* J_j *then the number of eigenvalues in the interval* (τ_1, τ_2) *can be obtained from the inertia of* $T(\tau_1)$ *and* $T(\tau_2)$ *according to Theorem 3.3.*

4. Quadratic eigenvalue problems.

We consider quadratic matrix pencils

$$Q(\lambda) := \lambda^2 A + \lambda B + C, \tag{4.1}$$

with Hermitian matrices $A, B, C \in \mathbb{C}^{n\times n}$ unter several conditions that guarantee that (some of) its real eigenvalues allow for a variational characterization, and hence for slicing of its spectrum using the inertia.

4.1. C < 0 and A ≥ 0.

Let C be negative definite and A positive semidefinite. Multiplying $Q(\lambda)x = 0$ by λ^{-1} one gets the equivalent nonlinear eigenvalue problem

$$\tilde{Q}(\lambda)x := \lambda A x + B x + \lambda^{-1} C x = 0. \tag{4.2}$$

Differentiating $f(\lambda; x) := x^H \tilde{Q}(\lambda) x$ with respect to λ yields

$$\frac{\partial}{\partial\lambda} f(\lambda; x) = x^H A x - \lambda^{-2} x^H C x > 0 \quad \text{for every } x \neq 0 \text{ and every } \lambda \neq 0.$$

Hence, $\tilde{Q}$ satisfies the conditions of the minmax characterization for both intervals $J_- := (-\infty, 0)$ and $J_+ := (0, \infty)$.

For the corresponding Rayleigh functional $p_\pm$ with domain $\mathcal{D}_\pm$ it holds that $\lambda_1^+ = \inf_{x\in\mathcal{D}_+} p_+(x) \in J_+$ and $\lambda_n^- = \sup_{x\in\mathcal{D}_i} p_-(x) \in J_-$, and therefore it follows from Theorem 3.2

THEOREM 4.1. *Let* C *be negative definite and* A *positive semidefinite.*

(i) For $\sigma > 0$ *let* $In(\tilde{Q}) = (n_p, n_n, n_z)$ *be the inertia of* $\tilde{Q}$*. Then the pencil* (4.1) *has* n_p *positive eigenvalues less than* σ*.*

(ii) For $\sigma < 0$ *let* $In(\tilde{Q}) = (n_p, n_n, n_z)$ *be the inertia of* $\tilde{Q}$*. Then* (4.1) *has* n_n *negative eigenvalues exceeding* σ*.*

If A is positive definite then $\tilde{Q}$ is overdamped with respect to J_+ and J_-, and therefore there exist exactly n positive and n negative eigenvalues.

If A is positive semidefinite and $r = \text{rank}(A)$, then ∞ is an infinite eigenvalue of multiplicity r, and there are only $2n - r$ finite eigenvalues. If B is positive definite, then the Rayleigh functional

$$p_+(x) = -2 \frac{x^H C x}{x^H B x + \sqrt{(x^H B x)^2 - 4(x^H A x)(x^H C x)}} \tag{4.3}$$

is defined on $\mathbb{C}^n \setminus \{0\}$. Hence, $(\tilde{Q}, J_+)$ is overdamped, and there exist n positive and $n - r$ negative eigenvalues.

4.2. Hyperbolic problems. The quadratic pencil $Q(\cdot)$ with Hermitian matrices $A, B, C \in \mathbb{C}^{n\times n}$ is called hyperbolic if A is positive definite and for every $x \in \mathbb{C}^n$, $x \neq 0$ the quadratic polynomial

$$f(\lambda; x) := \lambda^2 x^H Ax + \lambda x^H Bx + x^H Cx = 0$$

has two distinct real roots

$$p_\pm(x) = -\frac{x^H Bx}{2x^H Ax} \pm \sqrt{\left(\frac{x^H Bx}{2x^H Ax}\right)^2 - \frac{x^H Cx}{x^H Ax}}. \tag{4.4}$$

A hyperbolic quadratic matrix polynomial $Q(\cdot)$ has the following properties (cf. [12]): the ranges $J_\pm := p_\pm(\mathbb{C}^n \setminus \{0\})$ are disjoint real intervals with $\max J_- < \min J_+$, $Q(\lambda)$ is positive definite for $\lambda < \min J_-$ and $\lambda > \max J_+$, and it is negative definite for $\lambda \in (\max J_-, \min J_+)$.

Let $\tilde{J}_+$ be an open interval with $J_+ \subset \tilde{J}_+$ and $J_- \cap \tilde{J}_+ = \emptyset$, and let $\tilde{J}_-$ be an open interval with $J_- \subset \tilde{J}_-$ and $J_+ \cap \tilde{J}_- = \emptyset$. Then $(Q, \tilde{J}_+)$ and $(-Q, \tilde{J}_-)$ satisfy the conditions of the variational characterization of eigenvalues and they are both overdamped. Hence, there exist $2n$ eigenvalues

$$\lambda_1 \le \cdots \le \lambda_n < \lambda_{n+1} \le \cdots \le \lambda_{2n}$$

and

$$\lambda_j = \min_{\dim V = j} \max_{x\in V, x\neq 0} p_-(x) \quad \text{and} \quad \lambda_{n+j} = \min_{\dim V = j} \max_{x\in V, x\neq 0} p_+(x), \quad j = 1, \dots, n.$$

If $\mathrm{In}(Q(\sigma)) = (n_p, n_n, n_z)$ is the inertia of $Q(\sigma)$ and $n_n = n$, then $Q(\sigma)$ is negative definite and there are n eigenvalues smaller than σ and n eigenvalues exceeding σ. If $n_p = n$ then $Q(\sigma)$ is positive definite. $f(\sigma; x) > 0$ for every $x \neq 0$, and if $\frac{\partial}{\partial\lambda} f(\sigma; x) < 0$, then $\sigma < \lambda_1$, and otherwise $\sigma > \lambda_{2n}$.

If $n_p \neq n$ and $n_n \neq n$, then $\sigma \in J_- \cup J_+$ and Theorem 3.1 applies. We only have to find out in which of these intervals σ is located. To this end we determine $x \neq 0$ such that $f(\sigma; x) := x^H Q(\sigma)x > 0$ (this can be done by a few steps of the Lanczos method which is known to converge first to extreme eigenvalues). If $\frac{\partial}{\partial\lambda} f(\sigma; x) = 2\sigma x^H Ax + x^H Bx < 0$, then it follows that $p_-(x) > \sigma$, and therefore $\sigma \le \lambda_n = \max_{x\neq 0} p_-(x)$. Similarly, $f(\sigma; x) > 0$ and $2\sigma x^H Ax + x^H Bx > 0$ implies $\sigma > \lambda_{n+1} = \min_{x\neq 0} p_+(x)$. Hence we obtain the following slicing of the spectrum of $Q(\cdot)$.

THEOREM 4.2. *Let*

$$Q(\lambda) := \lambda^2 A + \lambda B + C$$

be hyperbolic, and let (n_p, n_n, n_z) *be the inertia of* $Q(\sigma)$ *for* $\sigma \in \mathbb{R}$.

- *(i) If* $n_n = n$ *then there are* n *eigenvalues smaller than* σ *and* n *eigenvalues greater than* σ.
- *(ii) Let* $n_p = n$. *If* $2\sigma x^H Ax + x^H Bx < 0$ *for an arbitrary* $x \neq 0$, *then there are* $2n$ *eigenvalues exceeding* σ, *if* $2\sigma x^H Ax + x^H Bx > 0$ *for an arbitrary* $x \neq 0$, *then all* $2n$ *eigenvalues are less than* σ.
- *(iii) For* $n_p = 0$ *and* $n_z > 0$ *let* $x \neq 0$ *be an element of the null space of* $Q(\sigma)$. *If* $2\sigma x^H Ax + x^H Bx < 0$ *then* $Q(\lambda)x = 0$ *has* $n - n_z$ *eigenvalues in* $(-\infty, \sigma)$, n *eigenvalues in* (σ, ∞) *and* $\sigma = \lambda_n$ *with multiplicity* n_z, *and if* $2\sigma x^H Ax + x^H Bx > 0$ *then* $Q(\lambda)x = 0$ *has* n *eigenvalues in* $(-\infty, \sigma)$, $n - n_z$ *eigenvalues in* (σ, ∞) *and* $\sigma = \lambda_{n+1}$ *with multiplicity* n_z.

(iv) For $n_p > 0$ and $n_z = 0$ let $x \neq 0$ such that $f(\sigma; x) > 0$. If $2\sigma x^H Ax + x^H Bx < 0$ then $Q(\lambda)x = 0$ has $n - n_p$ eigenvalues in $(-\infty, \sigma)$ and $n + n_p$ eigenvalues in (σ, ∞), and if $2\sigma x^H Ax + x^H Bx > 0$ then $Q(\lambda)x = 0$ has $n + n_p$ eigenvalues in $(-\infty, \sigma)$ and $n - n_p$ eigenvalues in (σ, ∞).

(v) For $n_p > 0$ and $n_z > 0$ let $x \neq 0$ such that $f(\sigma; x) > 0$. If $2\sigma x^H Ax + x^H Bx < 0$ then $Q(\lambda)x = 0$ has $n - n_p - n_z$ eigenvalues in $(-\infty, \sigma)$ and $n + n_p$ eigenvalues in (σ, ∞), and if $2\sigma x^H Ax + x^H Bx > 0$ then $Q(\lambda)x = 0$ has $n + n_p$ eigenvalues in $(-\infty, \sigma)$ and $n - n_p - n_z$ eigenvalues in (σ, ∞). In either case σ is an eigenvalue with multiplicity n_z.

REMARK 4.3. *These results on quadratic hyperbolic pencils can be generalized to hyperbolic matric polynomials of higher degree*

$$P(\lambda) = \sum_{j=0}^{k} \lambda^j A_j, \quad A_j = A_j^H, \ j = 0, \dots, k, \quad A_k > 0,$$

which is hyperbolic, if the polynomial $f(\lambda; x) := x^H P(\lambda)x = 0$ has exactly k real roots for every $x \neq 0$

Then there exist k disjoint open intervals $\Delta_j \subset \mathbb{R}$, $j = 1, \dots, k$ such that $P(\lambda)x = 0$ has exactly n eigenvalues in each Δ_j, and these eigenvalues allow for a minmax characterization. To fix the numeration let $\sup \Delta_{j+1} < \inf \Delta_j$ for $j = 1, \dots, k-1$.

For $\sigma \in \mathbb{R}$ let (n_p, n_n, n_z) be the inertia of $P(\sigma)$, and let $x \in \mathbb{C}^n$ such that $x^H P(\sigma)x > 0$. If $f(\cdot; x)$ has exactly j roots which exceed σ then it holds that

$$\sigma \in \Delta_{j+1} \quad \text{or} \quad \sigma \in [\sup \Delta_{j+1}, \inf \Delta_j] \quad \text{or} \quad \sigma \in \Delta_j.$$

Which one of these situations occur can be deduced from the inertia ($n_n = n$ or $n_p = n$) and the derivative $\frac{\partial}{\partial \lambda} f(\sigma; x)$ similar the quadratic case.

4.3. Definite quadratic pencils.

In a recent paper Higham, Mackey and Tisseur [10] generalized the concept of hyperbolic quadratic polynomials waiving the positive definiteness of the leading matrix A.

A quadratic pencil (4.1) is definite if $A = A^H$, $B = B^H$, and $C = C^H$ are Hermitian, there exists $\mu \in \mathbb{R} \cup \{\infty\}$ such that $Q(\mu)$ is positive definite, and for every $x \in \mathbb{C}^n$, $x \neq 0$ the quadratic polynomial

$$f(\lambda; x) := \lambda^2 x^H Ax + \lambda x^H Bx + x^H Cx = 0$$

has two distinct roots in $\mathbb{R} \cup \{\infty\}$.

The following Theorem was proved in [10].

THEOREM 4.4. *The Hermitian matrix polynomial $Q(\lambda)$ is definite if and only if any two (and hence all) of the following properties hold:*

— *$d(x) := (x^H Bx)^2 - 4(x^H Ax)(x^H Cx) > 0$ for every $x \in \mathbb{C}^n \setminus \{0\}$*

— *$Q(\eta) > 0$ for some $\eta \in \mathbb{R} \cup \{\infty\}$*

— *$Q(\xi) < 0$ for some $\xi \in \mathbb{R} \cup \{\infty\}$*

For $\xi < \eta$ (otherwise consider $Q(-\lambda)$) it was shown in [14] that there are n eigenvalues in (ξ, η) which are minmax values of a Rayleigh functional, and the remaining n eigenvalues in $[-\infty, \xi)$ and $(\eta, \infty]$ are maxmin and minmax values of a second Rayleigh functional. Hence, if ξ and η are known then the slicing of the spectrum using the LDL^H factorization follows similarly to the hyperbolic case. However, for a given σ and the LDL^H factorization of $Q(\sigma)$ there is no easy way to decide in which of the intervals $[-\infty, \xi)$, (ξ, η) or $(\eta, \infty]$ the parameter σ is located. [7, 8, 14] contain methods to detect whether a quadratic pencil is definite and to compute the parameters ξ and η, however they are much more costly than computing an LDL^H factorization of a matrix. For the Examples 4.5 and 4.7 at least one of these parameters are known and the slicing can be given explicitely.

EXAMPLE 4.5. *Duffin [4] called a quadratic eigenproblem* (4.1) *overdamped network, if A, B and C are positive semidefinite, and the so called overdamping condition*

$$d(x) := (x^H Bx)^2 - 4(x^H Ax)(x^H Cx) > 0 \quad \textit{for every } x > 0 \tag{4.5}$$

is satisfied.

So, actually B has to be positive definite, and therefore $Q(\mu)$ is positive definite for every $\mu > 0$, and $Q(\cdot)$ is definite.

If r denotes the rank of A, then (4.1) *has $n+r$ real eigenvalues the largest n ones of which (called primary eigenvalues by Duffin) are minmax values of*

$$p_+(x) := -2\frac{x^H Cx}{x^H Bx + \sqrt{d(x)}}$$

and the smallest r ones (called secondary eigenvalues) are maxmin values

$$\lambda_{n+1-j} = \max_{\dim V = j,\ V \cap \mathcal{D}_- \neq \emptyset} \ \min_{x \in V \cap \mathcal{D}_-} p_-(x),$$

where $\mathcal{D}_- := \{x \in \mathbb{C}^n : x^H Ax \neq 0\}$ and $p_-(x) = (-x^H Bx - \sqrt{d(x)})/(2x^H Ax)$ for $x \in \mathcal{D}_-$.

Hence, the following slicing of the spectrum follows

THEOREM 4.6. *Let $A, B, C \in \mathbb{C}^{n \times n}$ be positive semidefinite and assume that $d(x) > 0$ for $x \neq 0$. Let $In(Q(\sigma)) = (n_p, n_n, n_z)$ be the inertia of $Q(\sigma)$. Then it holds*

- *(i) If $n_n = r$ then there are r eigenvalues smaller than σ and n eigenvalues greater than σ.*
- *(ii) For $n_p > 0$ and $n_z = 0$ let $x \neq 0$ such that $f(\sigma; x) > 0$. If $2\sigma x^H Ax + x^H Bx < 0$ then $Q(\lambda)x = 0$ has $n-r-n_p$ eigenvalues in $(-\infty, \sigma)$ and $n+n_p$ eigenvalues in $(\sigma, 0)$, and if $2\sigma x^H Ax + x^H Bx > 0$ then $Q(\lambda)x = 0$ has $n-r+n_p$ eigenvalues in $(-\infty, \sigma)$ and $n-n_p$ eigenvalues in $(\sigma, 0)$.*

EXAMPLE 4.7. *Free vibrations of a fluid–solid structures are governed by the nonsymmetric eigenvalue problem [9, 18]*

$$\begin{bmatrix} K_s & C \\ 0 & K_f \end{bmatrix} \begin{bmatrix} x_s \\ x_f \end{bmatrix} = \lambda \begin{bmatrix} M_s & 0 \\ -C^T & M_f \end{bmatrix} \begin{bmatrix} x_s \\ x_f \end{bmatrix} \tag{4.6}$$

where $K_s \in \mathbb{R}^{s \times s}$ and $K_f \in \mathbb{R}^{f \times f}$ are the stiffness matrices, and $M_s \in \mathbb{R}^{s \times s}$ and $M_f \in \mathbb{R}^{f \times f}$ are the mass matrices of the structure and the fluid, respectively, and $C \in \mathbb{R}^{s \times f}$ describes the coupling of structure and fluid. x_s is the structure displacement vector, and x_f the fluid pressure vector. K_s, M_s, K_f and M_f are symmetric and positive definite.

Multplying the first line of (4.6) *by λ on obtains the quadratic pencil*

$$Q(\lambda) := \lambda^2 \begin{bmatrix} M_s & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} -K_s & -C \\ -C^T & M_f \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -K_f \end{bmatrix}. \tag{4.7}$$

It is easily seen that for $x_s \neq 0$ *the quadratic equation* $[x_s^T, x_f^T]Q(\lambda)\begin{bmatrix} x_s \\ x_f \end{bmatrix} = 0$ *has one positive solution* $p_+(x_s, x_f)$ *and one negative solution* $p_-(x_s, x_f)$*, and for* $x_s = 0$ *it has one positive solution* $p_+(x_s, x_f) := x_f^T K_f x_f / (x_f^T M_f x_f)$ *and the solution* $p_-(x_s, x_f) := \infty$*, and that the positive eigenvalues of* (4.6) *are minmax values of the Rayleigh functional* p_+*. Hence, one gets for the (physically meaningful) positive eigenvalues: If* $In(Q(\sigma)) = (n_p, n_n, n_z)$ *for* $\sigma > 0$*, then there are exactly* n_p *eigenvalues in* $(0, \sigma)$*,* n_n *eigenvalues in* (σ, ∞)*, and if* $n_z \neq 0$*, then* σ *is an eigenvalue of multiplicity* n_z*.*

4.4. Nonoverdamped quadratic pencils. We consider the quadratic pencil (4.1) where A, B and C are positive definite. Then for $x \neq 0$ the two complex roots of $f(\lambda; x) := x^H Q(\lambda) x$ are given in (4.4).

Let

$$\delta_- := \sup\{p_-(x) \ : \ \ p_-(x) \in \mathbb{R}\}, \quad \delta_+ := \inf\{p_+(x) \ : \ \ p_+(x) \in \mathbb{R}\}$$

and $J_- := (-\infty, \delta_+)$, $J_+ = (\delta_-, 0)$.

Then it is obvious that $-Q$ and Q satisfies the conditions of the minmax characterization of its eigenvalues in J_- and J_+, respectively. Hence, all eigenvalues in J_- are min max values of p_-

$$\lambda_j^- = \min_{\dim V = j,\ V \cap \mathcal{D}_- \neq \emptyset} \ \max_{x \in V \cap \mathcal{D}_-} \ p_-(x), \ \ j = 1, 2, \dots, \tag{4.8}$$

and all eigenvalues in J_+ are max min values of p_+

$$\lambda_{n+1-j}^+ = \max_{\dim V = j,\ V \cap \mathcal{D}_+ \neq \emptyset} \ \min_{x \in V \cap \mathcal{D}_+} \ p_+(x), \ \ j = 1, 2, \dots. \tag{4.9}$$

Hence, for $\sigma < \delta_+$ and for $\sigma > \delta-$ we obtain slicing results for the spectrum of $Q(\cdot)$ from Theorem 3.2. If $\mathrm{In}(Q(\sigma)) = (n_p, n_n, n_z)$ and $\sigma < \delta_+$, then there exist n_n eigenvalues of $Q(\cdot)$ in $(-\infty, \sigma)$, and if $\sigma \in (\delta_-, 0)$, then there are n_n eigenvalues in $(\sigma, 0)$. However, δ_+ and δ_- are usually not known. The following theorem contains upper bounds of δ_- and lower bounds of δ_+ thus yielding subintervals of $(-\infty, \delta_+)$ and $(\delta_-, 0)$ where the above slicing applies.

Theorem 4.8. *Let* $A, B, C \in \mathbb{C}^{n \times n}$ *positive definite, and let* p_+ *and* p_- *be defined in* (4.4)*. Then it holds that*

(i)

$$\tilde{\delta}_+ := -\sqrt{\max_{x \neq 0} \frac{x^H C x}{x^H A x}} \leq \delta_+ = \inf\{p_+(x) \ : \ \ p_+(x) \in \mathbb{R}\} \tag{4.10}$$

and

$$\delta_- = \sup\{p_-(x) \ : \ \ p_-(x) \in \mathbb{R}\} \leq -\sqrt{\min_{x \neq 0} \frac{x^H C x}{x^H A x}} =: \tilde{\delta}_-. \tag{4.11}$$

(ii)

$$\hat{\delta}_+ := -2 \max_{x \neq 0} \frac{x^H C x}{x^H B x} \leq \delta_+ < \delta_- \leq -2 \min_{x \neq 0} \frac{x^H B x}{x^H C x} =: \hat{\delta}_-. \tag{4.12}$$

Proof. (i): $\tilde\delta_+$ is a lower bound of δ_+, if for every $x \neq 0$ such that $p_+(x) \in \mathbb{R}$ and $\frac{\partial}{\partial\lambda} f(p_+(x); x) \geq 0$ it holds that $p_+(x) \geq \tilde\delta_+$.

$f(p_+(x); x) = x^H Q(p_+(x))x = 0$ if and only if

$$x^H Bx = -p_+(x)x^H Ax - \frac{1}{p_+(x)} x^H Cx.$$

Hence,

$$\frac{\partial}{\partial\lambda} f(p_+(x); x) = 2p_+(x)x^H Ax + x^H Bx = p_+(x)x^H Ax - \frac{1}{p_+(x)} x^H Cx \geq 0$$

if and only if

$$p_+(x)^2 \leq \frac{x^H Cx}{x^H Ax}, \text{ i.e. } \delta_+ \geq -\sqrt{\max_{x\neq 0} \frac{x^H Cx}{x^H Ax}} = \tilde\delta_+,$$

and analogously we obtain

$$\delta_- \leq -\sqrt{\min_{x\neq 0} \frac{x^H Cx}{x^H Ax}} = \tilde\delta_-.$$

(ii): Solving $f(p_+(x); x) = 0$ for $x^H Ax$ one gets from $\frac{\partial}{\partial\lambda} f(p_+(x); x) \geq 0$

$$p_+(x) \geq -2\frac{x^H Cx}{x^H Bx}, \text{ i.e. } \delta_+ \geq -2\max_{x\neq 0} \frac{x^H Cx}{x^H Bx} = \hat\delta_+,$$

and analogously

$$\delta_- \leq -2\min_{x\neq 0} \frac{x^H Cx}{x^H Bx} = \hat\delta_-.$$

□

From Theorem 3.2 we obtain the following slicing of the spectrum of $Q(\cdot)$.

THEOREM 4.9. *Let A, B and C be positive definite, and for $\sigma \in \mathbb{R}$ let $In(Q(\sigma)) = (n_p, n_n, n_z)$.*

(i) Let

$$\sigma \leq \max\left\{-\sqrt{\max_{x\neq 0} \frac{x^H Cx}{x^H Ax}}, -2\max_{x\neq 0} \frac{x^H Cx}{x^H Bx}\right\}.$$

Then there exist n_n eigenvalues of $Q(\lambda)x = 0$ in $(-\infty, \sigma)$.

(ii) Let

$$\sigma \geq \min\left\{-\sqrt{\min_{x\neq 0} \frac{x^H Cx}{x^H Ax}}, -2\min_{x\neq 0} \frac{x^H Cx}{x^H Bx}\right\}.$$

Then there exist n_n eigenvalues of $Q(\lambda)x = 0$ in $(\sigma, 0)$.

EXAMPLE 4.10. *The matrices A, B and C were obtained from the following MATLAB statements:*

```
randn('state',0); A=eye(20); B=randn(20);B=B'*B; C=randn(20);C=C'*C;
```

Then $Q(\lambda)x = 0$ *has* 26 *real eigenvalues,* 13 *in the domain of* p_- *and* 13 *in the domain of* p_+*. So, Sylvester's theorem can be applied to all of them.* 12 *eigenvalues are less than* $-\sqrt{\max(\lambda(C, A))}$ *and* 6 *eigenvalues exceed* $-\sqrt{\min(\lambda(C, A))}$*.*

REFERENCES

[1] C. Conca, J. Planchard, and M. Vanninathan, *Fluid and Periodic Structures*, vol. 24 of Research in Applied Mathematics, Masson, Paris, 1995.

[2] R. Courant, *Über die Eigenwerte bei den Differentialgleichungen der mathematischen Physik*, Math. Z., 7 (1920), pp. 1 – 57.

[3] J. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[4] R. Duffin, *A minimax theory for overdamped networks*, J. Rat. Mech. Anal., 4 (1955), pp. 221 – 233.

[5] E. Fischer, *Über quadratische Formen mit reellen Koeffizienten*, Monatshefte für Math. und Phys., 16 (1905), pp. 234 – 249.

[6] G. Golub and C. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore and London, 3rd ed., 1996.

[7] C.-H. Guo, N. Higham, and F. Tisseur, *Detecting and solving hyperbolic quadratic eigenvalue problems*, SIAM J. Matrix Anal. Appl., 30 (2009), pp. 1593 – 1613.

[8] ———, *An improved arc algorithm for detecting definite Hermitian pairs*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1131 – 1151.

[9] C. Harris and A. Piersol, eds., *Harris Shock and Vibration Handbook*, McGraw-Hill, New York, 5 ed., 2002.

[10] N. Higham, D. Mackey, and F. Tisseur, *Definite matrix polynomials and their linearization by definite pencils*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 478 – 502.

[11] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[12] A. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, vol. 71 of , Translations of Mathematical Monographs, AMS, Providence, 1988.

[13] A. Neumaier, *Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, 2001.

[14] V. Niendorf and H. Voss, *Detecting hyperbolic and definite matrix polynomials*, Linear Algebra Appl., 432 (2010), pp. 1017 – 1035.

[15] B. Parlett, *The symmetric eigenvalue problem*, SIAM, Philadelphia, 1998.

[16] H. Poincaré, *Sur les equations aux dérivées partielles de la physique mathématique*, Amer. J. Math., 12 (1890), pp. 211 – 294.

[17] E. Rogers, *A minimax theory for overdamped systems*, Arch. Ration. Mech. Anal., 16 (1964), pp. 89 – 96.

[18] M. Stammberger and H. Voss, *On an unsymmetric eigenvalue problem governing free vibrations of fluid-solid structures*, Electr. Trans. Numer. Anal., 36 (2010), pp. 113 – 125.

[19] J. Sylvester, *A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by orthogonal substitutionsto the form of a sum of positive and negative squares*, Philosophical Magazin, 4(23) (1852), pp. 138 – 142.

[20] H. Voss, *A minmax principle for nonlinear eigenproblems depending continuously on the eigenparameter*, Numer. Lin. Algebra Appl., 16 (2009), pp. 899 – 913.

[21] H. Voss and B. Werner, *A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems*, Math. Meth. Appl. Sci., 4 (1982), pp. 415–424.