

Inertia laws and localization of real eigenvalues for generalized indefinite eigenvalue problems

Yuji Nakatsukasa^a, Vanni Noferini^{b,*}

^a *Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom*

^b *Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland*

Abstract

Sylvester's law of inertia states that the number of positive, negative and zero eigenvalues of Hermitian matrices is preserved under congruence transformations. The same is true of generalized Hermitian definite eigenvalue problems, in which the two matrices are allowed to undergo different congruence transformations, but not for the indefinite case. In this paper we investigate the possible change in inertia under congruence for generalized Hermitian indefinite eigenproblems, and derive sharp bounds that show the inertia of the two individual matrices often still provides useful information about the eigenvalues of the pencil, especially when one of the matrices is almost definite. A prominent application of the original Sylvester's law is in finding the number of eigenvalues in an interval. Our results can be used for estimating the number of real eigenvalues in an interval for generalized indefinite and nonlinear eigenvalue problems.

Keywords: Sylvester's law of inertia, generalized indefinite eigenvalue problem, number of eigenvalues in an interval, congruence transformation, nonlinear eigenvalue problems
2000 MSC: 15A18, 15A22, 65F15

1. Introduction

Let $A = A^* \in \mathbb{C}^{n \times n}$. The *inertia* [13] of A is the triple

$$(n_+(A), n_0(A), n_-(A)) \in \mathbb{N}^3 \quad (1)$$

where $n_+(A)$, $n_0(A)$, and $n_-(A)$ denote, respectively, the number of positive, zero, and negative eigenvalues of A . Clearly, the inertia cannot be any triple of nonnegative integers, because its elements must satisfy the constraints

$$n_+(A) + n_0(A) + n_-(A) = n, \quad n_+(A) + n_-(A) = \text{rank}(A).$$

It is also common to define the *signature* of A as $s(A) = n_+(A) - n_-(A)$.

*Corresponding author.

Email addresses: `nakatsukasa@maths.ox.ac.uk` (Yuji Nakatsukasa), `vanni.noferini@aalto.fi` (Vanni Noferini)

Preprint submitted to Linear Algebra and its Applications

May 17, 2019

Sylvester's law of inertia [13] identifies the orbits by congruence as the equivalence classes prescribed by inertia. It is named after J. J. Sylvester, who first proved the result [33]. We give below a formal statement and a concise (albeit not elementary) algebraic proof.

Theorem 1 (Sylvester's law of inertia) *Let $A = A^*, B = B^* \in \mathbb{C}^{n \times n}$. A and B are congruent, i.e., there exists an invertible matrix $X \in \mathbb{C}^{n \times n}$ such that $A = X^*BX$, if and only if A and B have the same inertia.*

Proof. Suppose that A and B have the same inertia. Then, with no loss of generality we may assume that A and B are diagonal (by the spectral theorem) and that A_{ii} and B_{ii} have the same sign (if not, note that there exist a permutation matrix P such that $B = P^T C P$ and A, C have this property). Hence, one can define X to be diagonal with $X_{ii} = (A_{ii}/B_{ii})^{1/2}$ if $B_{ii} \neq 0$ and $X_{ii} = 1$ otherwise.

For the converse, observe that $n_+(A)$ is the maximal dimension of a subspace over which the quadratic form defined by A is positive definite; as such, it is invariant by any change of basis (which is equivalent to a congruence on A). Similarly, the rank is invariant by congruence. We conclude that if A and B are congruent then they must have the same inertia. \square

While the “if” direction of this proof is not much more than an immediate corollary of the spectral theorem, the argument to prove the “only if” implication is arguably more advanced, and it implicitly relies on the minmax Rayleigh characterization of the eigenvalues of a Hermitian matrix.

Sylvester's law of inertia is a useful tool in many applications, including counting the number of eigenvalues of a symmetric matrix in an interval and the design of algorithms that compute eigenvalues (or singular values) with high relative accuracy [9, Ch. 5], counting the number of eigenvalues above (or below) a certain value.

A few generalizations of Sylvester's law of inertia are available in the literature. Kostić and Voss [17] introduced Sylvester-like laws of inertia for nonlinear eigenvalue problems, but they dealt with situations in which a minmax-type characterization of the eigenvalues exist. Instead, we are mainly concerned with those that do not fall into this category, for example the generalized indefinite eigenvalue problem, so that the results in [17] are inapplicable. Bilir and Chicone [3] studied the inertia of the real parts of the eigenvalues of quadratic matrix polynomials whose leading and trailing coefficients are Hermitian, and the second coefficient has a positive definite Hermitian part. Going back to matrix, as opposed to nonlinear, eigenvalue problems, Ikramov [14] showed that two normal matrices are congruent if and only if they have the same number of eigenvalues on any semiline starting from 0 in the complex plane.

In this work, we aim to extend Sylvester's classical result in a different direction. Specifically, we will analyze generalized Hermitian (*indefinite*, that is, the matrix $\alpha A + \beta B$ is never definite for any $(\alpha, \beta) \in \mathbb{R}^2$), polynomial, and generally nonlinear eigenvalue problems. We will see that, although it is not generally possible to determine the number of positive, zero, negative, and overall real eigenvalues, there exist nontrivial bounds on these quantities that only depend on the inertia of the trailing and (when appropriate) leading coefficients of the eigenvalue problem, and as such are invariant by congruence on them.

As in the original law of inertia, our approach can be used to obtain bounds for the number of eigenvalues lying in an interval for generalized Hermitian indefinite eigenvalue problems, which are particularly useful when one of the matrices is “nearly” positive (or negative) definite, by which we mean that most of the eigenvalues are positive (or negative). Moreover, we show that such results can be extended to nonlinear eigenvalue problems.

Our results are useful in a number of applications. For example, estimating the number of eigenvalues in an interval is a key component for an efficient eigensolver based on splitting the spectrum, for example using contour integration; see [10] and the references therein. Generalized Hermitian indefinite eigenvalue problems arise for instance in optimization problems [1], tensor decomposition [7], and when one uses a structured-preserving linearization of a Hermitian matrix polynomial (e.g., one arising from a bivariate zero-finding problem [27]), for example the \mathbb{DL} linearization [22, 28] or the family of block symmetric linearizations described in [6]. In all these applications [1], [7] and [27], it is the real eigenvalues that are of interest. Eigenvalues of indefinite pairs play a crucial role also in preconditioning, for example when a constraint preconditioner is used for indefinite linear systems [16].

The structure of the paper is as follows. In Section 2 we revisit Sylvester’s law of inertia for classical and generalized Hermitian definite eigenvalue problems. In particular, we give an analytic proof of the “only if” implication in Theorem 1 by using elementary arguments, namely continuity of eigenvalues and the intermediate value theorem. In Section 3 we extend the argument to analyze the generalized indefinite eigenvalue problem and obtain bounds for its inertia. In Section 4 we discuss applications and variants, including bounding the number of eigenvalues in any real interval, i.e., not necessarily $(0, \infty)$ or $(-\infty, 0)$. In Section 5 we discuss the polynomial eigenvalue problem, and more generally the nonlinear eigenvalue problem. Until Section 5 we focus on counting eigenvalues with their algebraic multiplicity; Section 6 discusses bounds when counting eigenvalues using their geometric multiplicity. Numerical experiments are presented in Section 7 to illustrate the results.

2. Sylvester’s law of inertia revisited

The following result (see [15, Sec. II.5] and [25, Sec. 5] for a more thorough discussion) is key for our analysis.

Theorem 2 *Let $\Omega \subseteq \mathbb{R}$ be an interval, and let $F(t)$ be an $n \times n$ complex matrix whose entries depend continuously on a real parameter t and such that for all $t \in \Omega$ the eigenvalues of $F(t)$ are real. Then, there exist n continuous functions $\lambda_i(t)$, $i = 1, \dots, n$, such that for all $t \in \Omega$ they are the eigenvalues of $F(t)$.*

In this paper, we will take $F(t)$ to be Hermitian for all values of $t \in \Omega$, which of course guarantees that its eigenvalues are real. F. Rellich [31, 32] pioneered the study of this special case, and we will sometimes call the $\lambda_i(t)$ “Rellich’s eigenfunctions”, or just “eigenfunctions” for brevity¹.

¹In other contexts, “eigenfunction” may mean the eigenvector of linear operators defined on functional spaces. We never use the term in that sense in this paper, so no ambiguity should arise. Eigenfunctions in this sense are also a classical and important object in operator theory [24, Ch. IV].

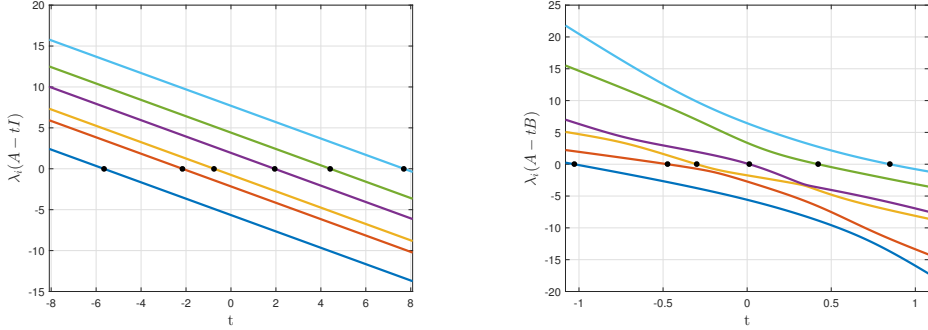


Figure 1: Eigenvalues of 6×6 matrices of form $A - tI$ (left) and $A - tB$, $B \succ 0$ (right). The eigenvalues of (A, B) , shown as the abscissae of the black dots, are the roots of the curves. The inertiae of A , and hence of the pencil, are $(3, 0, 3)$ in both figures.

We first set $F(t) = A - tI$ where $A = A^* \in \mathbb{C}^{n \times n}$ has inertia $(n_+(A), n_0(A), n_-(A))$. Clearly, the assumptions of Theorem 2 are satisfied on $\Omega = \mathbb{R}$, and it is particularly easy to write down explicitly the functions

$$\lambda_i(t) = \lambda_i(A) - t,$$

where $\lambda_i(A)$ is the i th eigenvalue of A . With this view, the eigenvalues of A are precisely the values of t for which one of Rellich's eigenfunctions $\lambda_i(t)$ of $F(t)$ has a zero. See Figure 1 (left) for an illustration.

Our second step is to consider a generalized Hermitian definite eigenvalue problem, and to study its eigenvalues (whose formal definition is given in Section 3). This corresponds to $F(t) = A - tB$, with $B = B^* \neq I$ but $B \succ 0$. The eigenfunctions $\lambda_i(t)$ are no longer necessarily affine functions, but they are still strictly decreasing functions. This can be proved, for instance, by Weyl's theorem. See Figure 1 (right). In particular, for this choice of $F(t)$ one has that $\lambda_i(t) \rightarrow -\infty$ as $t \rightarrow +\infty$ and $\lambda_i(t) \rightarrow +\infty$ as $t \rightarrow -\infty$; implying that if a curve $\lambda_i(t)$ takes a positive value (resp. negative) at $t = 0$, then it must have a root on $(0, \infty)$ (resp. $(-\infty, 0)$). Therefore the number of positive (resp., zero, negative) eigenvalues is still exactly equal to that of A . This observation constitutes an analytic proof of the “only if” part of Theorem 1. Indeed, it suffices to observe that $\det(X^*AX - tI) = \det(X)^2 \det(A - tX^{-*}X^{-1})$ and that $X^{-*}X^{-1}$ is positive definite by construction.

So far, we have provided an unusual analytic approach to the subject of inertia to reinvent the wheel, but we have not obtained any new results. Crucially, though, this technique relies on the very general Theorem 2, and therefore it is particularly suitable to be carried over to other types of eigenvalue problems. Before proceeding with this task, let us gain further insight into the difficulties that arise in locating the eigenvalues of the pencil when the definiteness assumption is dropped by setting $F(t) = A - tB$ with A and B Hermitian, but B no longer positive definite². Proofs analogous to the one we gave for Theorem 1, relying on the minmax characterization, will clearly face the difficulty that such characterizations are unavailable.

²Note that the cases of B negative definite or A (positive or negative) definite are also trivial, in the sense of being analogous to the the case $B \succ 0$. For example, we can set, respectively, $F(t) = A + tB$,

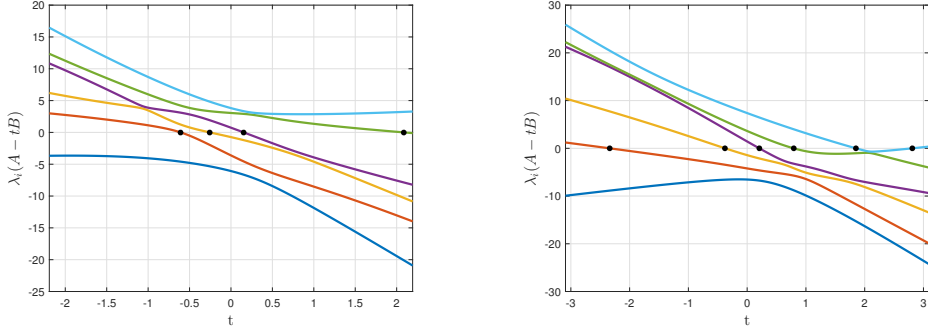


Figure 2: Plot of the eigenvalues of $A - tB$ for A, B both indefinite. The inertia of A is $(3, 0, 3)$ in both cases. Left: B has inertia $(5, 0, 1)$. Our result proves that (A, B) must have at least two negative eigenvalues. Right: B also has inertia $(3, 0, 3)$. Our result gives a lower bound of zero negative eigenvalues, but there are two of them.

To illustrate the idea, suppose that B has only 1 negative eigenvalue and $n - 1$ positive eigenvalues. Then, our results show that at least $n - 2$ eigenvalues are real; see the left plot of Figure 2 for an illustration with $n = 6$. The inertia of A is $(3, 0, 3)$, so three of the six curves $\lambda_i(A - tB)$ take positive values at $t = 0$. As $t \rightarrow \infty$, however, the inertia of $A - tB$ must match that of $-B$, which is $(1, 0, 5)$. Hence, five curves must be negative for t sufficiently large. It follows from the intermediate value theorem that at least two curves must intersect the t -axis. We conclude that at least two of the generalized eigenvalues of the pencil (A, B) are positive, and analogously the pencil has at least two negative eigenvalues. In this example the bound matches the exact value, but to be convinced that such lower bound may be an underestimate, see the right plot of Figure 2. We have kept the same A , but changed B so that it now has inertia $(3, 0, 3)$. The intermediate value theorem gives the trivial lower bound zero for the number of negative eigenvalues. Nonetheless, (A, B) has two negative eigenvalues.

In Section 3, we make these observations more general and more precise.

3. Inertia bounds for indefinite pencils

Given $A = A^*, B = B^* \in \mathbb{C}^{n \times n}$, the Hermitian generalized eigenvalue problem is to find the values of z such that the rank of $A - zB$ is not maximal. Formally, we define $A - \infty B := B$ and $\mathbb{C}^* := \mathbb{C} \cup \{\infty\}$. We say that $\lambda \in \mathbb{C}^*$ is an eigenvalue of the pencil $A - zB$, or equivalently of (A, B) , if

$$\text{rank}(A - \lambda B) < \max_{z \in \mathbb{C}^*} \text{rank}(A - zB) =: r. \quad (2)$$

This generalizes the standard definition $Ax = \lambda Bx$, $0 \neq x \in \mathbb{C}^n$, which is valid for *regular* generalized eigenvalue problems, i.e., when $r = n$. The right-hand side of the

$F(t) = tA - B$, $F(t) = tA + B$. Similarly, if a linear combination $\alpha A + \beta B$ for some $\alpha, \beta \in \mathbb{R}$ is positive definite, then one can consider $F(t) = A - t(\alpha A + \beta B)$: indeed, note that the eigenfunctions of the latter have a zero at $\frac{\lambda}{\alpha + \beta \lambda}$ if and only if λ is an eigenvalue of the generalized eigenvalue problem $A - zB$.

above equation, r , is sometimes called the *normal rank*, or simply rank, of the pencil $A - zB$. Note that, invoking Theorem 2, this definition implies that a real eigenvalue must be a zero of at least one non-identically zero eigenfunction $\lambda_i(t)$ of $F(t) = A - tB$, $t \in \mathbb{R}$. Definition (2) also ensures that there are at most r eigenvalues, counted with (either algebraic or geometric) multiplicities [8]. Here, the geometric multiplicity of an eigenvalue λ is the rank drop $r - \text{rank}(A - \lambda B)$. The algebraic multiplicity is the sum of the multiplicities of the zeros of the non-identically zero eigenfunctions; see [25, Def. 2.3]. Clearly, the algebraic multiplicity is at least as large as the geometric counterpart.

There may be strictly less than r eigenvalues; for example, if

$$A - zB = \begin{bmatrix} 0 & 1 & -z \\ 1 & 0 & 0 \\ -z & 0 & 0 \end{bmatrix} \quad (3)$$

there are no eigenvalues, even though the rank is 2. See, e.g., [8] for a sharp characterization of the number of the eigenvalues.

As we have seen in Section 2, when B (or A) is positive (or negative) definite, the number of positive, zero, and negative eigenvalues can be determined from the inertia of A (or B), and this task has a close relationship with Sylvester's classical result on the inertia of matrices. Similar remarks can be made when there exist scalars $\alpha, \beta \in \mathbb{R}$ such that $\alpha A + \beta B$ is positive definite; we will discuss this situation in detail in Section 4.

When no such (α, β) can be found, the generalized eigenvalue problem is said to be indefinite, and it becomes somewhat more difficult. One complication with generalized indefinite eigenvalue problems is the potential existence of nonreal eigenvalues and eigenvalues at infinity, defined as eigenvalues at 0 of $B - zA$. For example, the pencil

$$A - zB = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & z \\ 0 & z & 1 \end{bmatrix} \quad (4)$$

is indefinite, and indeed its eigenvalues are ∞, i , and $-i$. Another complication is that the pencil may be singular, i.e., the normal rank is deficient, or equivalently, that $\det(A - zB) = 0$ for all $z \in \mathbb{C}$. For example, the pencil (3) is singular.

Remark 3 *Having a connection with Sylvester's law of inertia in mind, we also note that it may happen that $A - zB$ is a regular pencil, i.e. $\det(A - zB)$ is not identically zero, but there exist invertible X, Y for which the pencil $X^*AX - zY^*BY$ is singular. For example, take*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad X = I_2, \quad Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (5)$$

*The results that we develop in the current section provide localization bounds for the eigenvalues that only depend on the inertia of the two matrix coefficients of the generalized eigenvalue problem, and therefore are invariant under such congruence transformations of the pencil. In other words, we give bounds that hold regardless of whether or not the resulting pencil $X^*AX - zY^*BY$ is singular.*

In order to take into account eigenvalues that are nonreal and at infinity, we define the inertia of $A - zB$ as the quintuple

$$(n_+(A, B), n_0(A, B), n_-(A, B), n_C(A, B), n_\infty(A, B)). \quad (6)$$

The entries of the inertia of $A - zB$ are nonnegative integers that count the number of, resp., positive, zero, negative, nonreal finite, and infinite eigenvalues, where each eigenvalue is counted with its algebraic multiplicity.

Remark 4 *Observe that the inertiae of $A - zB$ and $B - zA$ are related by*

$$\begin{aligned} & (n_+(A, B), n_0(A, B), n_-(A, B), n_C(A, B), n_\infty(A, B)) \\ &= (n_+(B, A), n_\infty(B, A), n_-(B, A), n_C(B, A), n_0(B, A)). \end{aligned}$$

Most, but not all, bounds we derive are also valid for the geometric multiplicity; we describe this in Section 6. We note that our definition is different from the one by Bilir and Chicone [3], who examined the real parts of the eigenvalues (and hence no $n_C(A, B)$ arises).

There appear to be no known bounds for the inertia (6) of generalized indefinite eigenvalue problems. The purpose of Theorem 5 below is to fill in this gap.

Theorem 5 *Let A, B be Hermitian matrices and denote their inertia by $(n_+(A), n_0(A), n_-(A))$ and $(n_+(B), n_0(B), n_-(B))$. Also define for notational convenience*

$$\begin{aligned} n_{++} &= n_+(A) + n_+(B), & n_{--} &= n_-(A) + n_-(B), \\ n_{+-} &= n_+(A) + n_-(B), & n_{-+} &= n_-(A) + n_+(B), \\ \delta &= n_0(A) - n_0(B), \end{aligned}$$

and $N_{++} = \max(n_{++}, n_{--})$, $N_{+-} = \max(n_{+-}, n_{-+})$. Then the inertia of the generalized eigenvalue problem $A - zB$ as in (6) satisfies

$$N_{++} - n \leq n_+(A, B) \leq 2n - |\delta| - N_{+-}, \quad (7)$$

$$\delta \leq n_0(A, B) \leq 3n - N_{++} - N_{+-} - n_0(B), \quad (8)$$

$$N_{+-} - n \leq n_-(A, B) \leq 2n - |\delta| - N_{++}, \quad (9)$$

$$0 \leq n_C(A, B) \leq 2 \min(n_+(A), n_-(A), n_+(B), n_-(B)), \quad (10)$$

$$-\delta \leq n_\infty(A, B) \leq 3n - N_{++} - N_{+-} - n_0(A). \quad (11)$$

Moreover, denoting by $n_R(A, B)$ the number of finite real eigenvalues of the generalized eigenvalue problem $A - zB$, we have

$$|s(B)| \leq n_R(A, B) \leq n - n_0(B). \quad (12)$$

Remark 6 *Of course, we also have the trivial bounds: lower bound 0 and upper bound n . It is possible that some of the bounds (7)–(10) are no more informative.*

Proof. [Proof of Theorem 5] We start from zero and infinite eigenvalues. We need the following definitions: let r be the rank of the pencil $A - zB$. Let $N = n_+(A, B) + n_-(A, B) + n_0(A, B) + n_\infty(A, B) + n_C(A, B)$ be the total number of eigenvalues counted with algebraic multiplicities. Then $N \leq r$. The number of intersections of the Reillich's eigenfunctions with the real axis is equal to, counting multiplicities, $N - n_C(A, B) - n_\infty(A, B)$. Furthermore, there are $n - r$ eigenfunctions identically equal to zero.

The number of Rellich's eigenfunctions which are 0 at $t = 0$ is $n_0(A)$; of those, $n - r$ correspond to identically zero eigenfunctions. Hence, it is immediate that $n_0(A, B) \geq n_0(A) - (n - r)$. Using $\text{rank}(B) = n - n_0(B) \leq r$, we note that $n - r \leq n_0(B)$, and so

$$n_0(A, B) \geq n_0(A) - n_0(B).$$

Switching the roles of A and B yields the lower bound for $n_\infty(A, B)$.

Now we come to the positive eigenvalues. To avoid normalization issues at infinity, it will be convenient to consider the mapping $A - zB \mapsto \sin(\theta)(A - \cot(\theta)B)$ and look at the eigenvalues of $A \sin(\theta) - B \cos(\theta)$ for $\theta \in [0, \pi]$, and at the corresponding eigenfunctions as functions of θ . The number of positive eigenvalues is the number of zeros in $(0, \pi/2)$ of the non-identically zero eigenfunctions. Let us classify each eigenfunction according to its sign at $\theta = \pi/2$ ($z = 0$) and its sign at $\theta = 0$ ($z = +\infty$). Note that there are 9 possible types.

Any eigenfunction that is positive at $\theta = \pi/2$ corresponds to a positive eigenvalue of A ; in order for this eigenfunction not to correspond to a real positive eigenvalue of the pencil, then it must be ≥ 0 at $\theta = 0$, corresponding to a nonpositive eigenvalue of B . Hence, we have the lower bound

$$n_+(A, B) \geq n_+(A) - n_0(B) - n_-(B) = n_+(A) + n_+(B) - n.$$

Since an analogous argument can be given by switching the roles of $+$ and $-$, we also obtain

$$n_+(A, B) \geq n_-(B) + n_-(A) - n.$$

The lower bounds for the negative eigenvalues are obtained immediately via the mapping $B \mapsto -B$.

For the upper bounds, observe that the number of positive eigenvalues is bounded above by the total number of eigenvalues, minus the number of infinite, zero and negative eigenvalues. Using the bounds previously obtained this yields

$$n_+(A, B) \leq r - |\delta| + \min(n - n_-(A) - n_+(B), n - n_+(A) - n_-(B))$$

Observe now that $r \leq n$ to obtain the bound for positive eigenvalues. The bound for negative eigenvalues is obtained similarly by sending $B \mapsto -B$. The upper bounds for zero and infinite eigenvalues are obtained analogously, by using the lower bounds. We omit the details.

We turn to (12). For the upper bound, observe that, arguing as above but switching the roles of A and B , we obtain $n_\infty(A, B) \geq n_0(B) - (n - r)$. Hence,

$$n_R(A, B) \leq r - n_\infty(A, B) \leq n - n_0(B).$$

For the lower bound $s(B)$, observe that, setting $F(t) = A - tB$ in Theorem 2, the Rellich's eigenfunctions $\lambda_i(t)$ must (altogether) have at least $|n_+(B) - n_-(B)|$ zeros on $(-\infty, \infty)$.

It remains to establish the upper bound for the number of complex eigenvalues (10). We trivially have $n_C(A, B) \leq n - n_R(A, B) - n_\infty(A, B)$. Now since $n_\infty(A, B) \geq n_0(B) - (n - r)$, using (12) we obtain

$$n_C(A, B) \leq r - n_R(A, B) - n_\infty(A, B) \leq n - n_0(B) - |n_+(B) - n_-(B)|.$$

Consider the case $n_+(B) > n_-(B)$. Then since $n_+(B) = n - n_-(B) - n_0(B)$, we obtain

$$n_C(A, B) \leq n - n_0(B) - (n_+(B) - n_-(B)) = 2n_-(B).$$

The case $n_+(B) \leq n_-(B)$ is analogous. By symmetry of the statement (10) with respect to the matrices A, B , we also conclude that $n_C(A, B)$ is bounded by $2n_-(A)$ and $2n_+(A)$. \square

It is worth remarking that all the bounds that appear in Theorem 5 are invariant under congruence $A \mapsto X^*AX$, $B \mapsto Y^*BY$. Therefore all the bounds remain the same when applied to (A, B) or (X^*AX, Y^*BY) for nonsingular matrices X, Y . In particular, if $B \succ 0$ then B is congruent to I and we recover the classical Sylvester's theorem of inertia: the bounds (7)–(9) become equalities on $n_-(X^*AX, Y^*BY)$, $n_0(X^*AX, Y^*BY)$ and $n_+(X^*AX, Y^*BY)$. Otherwise (when B is not definite), the congruence transformation may result in different $n_-(X^*AX, Y^*BY)$, $n_0(X^*AX, Y^*BY)$ and $n_+(X^*AX, Y^*BY)$, but within the intervals given in (7)–(9).

Remark 7 *An alternative, and more algebraic, approach to proving Theorem 5 is possible via the canonical form of Hermitian pencils via congruence [18, 34]; see also [21, 26] for related work. We prefer the analytic approach, because it generalizes to polynomial and nonlinear eigenvalue problems, for which no canonical form is available.*

Theorem 8 *The lower bounds in (7)–(11) in Theorem 5 are sharp: given any Hermitian pair A, B and given any of the corresponding lower bounds in (7)–(11), suppose that the bound is not trivial (i.e., it is not below 0 or above r). Then, there exist nonsingular matrices X, Y such that for the pencil $X^*AX - zY^*BY$, equality is attained for the bound considered. In particular, it is impossible to obtain a nontrivial lower bound for $n_C(A, B)$ from the inertia information of A, B .*

Proof. The proof is constructive: for each lower bound, we exhibit a Hermitian pair (\hat{A}, \hat{B}) such that A , resp. B , has the same inertia as \hat{A} , resp. \hat{B} , and the pencil $\hat{A} - z\hat{B}$ achieves the bound. The result will then follow by Theorem 1.

- *Equations (7), (9), (10).* Let us first show how the lower bound on positive eigenvalues can be attained. Without loss of generality, we take $n_{++} \geq n_{--}$ and $n_+(A) \geq n_-(B)$. Indeed, note that $A - zB$, $zB - A$, $B - zA$, and $zA - B$ all have the same number of positive eigenvalues, so if necessary we can swap the roles of A and B and/or multiply both matrices times -1 . We take

$$\hat{A} = \begin{bmatrix} I_{n_+(A)} & 0 & 0 \\ 0 & -I_{n_-(A)} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} -I_{n_-(B)} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{n_+(B)} \end{bmatrix}.$$

9

By the assumptions above, the only possibility for a positive eigenvalue is if the blocks $I_{n_+(A)}$ and $I_{n_+(B)}$ superpose, which either happens precisely on a sub-block of size $n_{++} - n$ (if this is nonnegative), or it does not happen at all otherwise. Note that this example also attains the lower bound of 0 complex eigenvalues, and that again $A - zB$, $zB - A$, $B - zA$, and $zA - B$ all have the same number of complex eigenvalues.

For the lower bound on negative eigenvalues, we assume without loss of generality that $n_{+-} \geq n_{-+}$ and $n_+(A) \geq n_+(B)$, and we change the matrices from the example above as follows: $(\hat{A}, \hat{B}) \rightarrow (\hat{A}, -\hat{B})$.

- *Equations (8), (11).* We only discuss the lower bound for zero eigenvalues, as the case of infinite eigenvalues can be proved similarly switching the roles of A and B . We take

$$\hat{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{n_+(A)} & 0 \\ 0 & 0 & -I_{n_-(A)} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{n_+(B)} & 0 \\ 0 & 0 & -I_{n_-(B)} \end{bmatrix},$$

and we observe that $\hat{A} - z\hat{B}$ has no zero eigenvalues if $n_0(A) \leq n_0(B)$, and it has δ zero eigenvalues otherwise.

□

Remark 9 *The lower bound $|s(B)| \leq n_R(A, B)$ in (12) is sharp if we assume the knowledge of only $(n_+(B), n_0(B), n_-(B))$. For instance, take*

$$\hat{A} = \begin{bmatrix} R & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} T & 0 & 0 \\ 0 & \pm I_{|s(B)|} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where $p = \min(n_+(B), n_-(B))$, the sign in $\pm I_{|s(B)|}$ takes $+1$ if $n_+(B) > n_-(B)$ and -1 otherwise, D_1 and D_2 are any real diagonal matrices, and

$$R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathbb{C}^{2p \times 2p}, \quad T = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in \mathbb{C}^{2p \times 2p}. \quad (13)$$

Obtaining a sharp bound using the inertia of both A and B is complicated by the possible presence of eigenvalues at 0 and ∞ . We obtain the lower bound $n_R(A, B) \geq N_{++} + N_{+-} - 2n + \delta$ by adding (7)–(9), but neither this nor $|s(B)|$ is always sharp.

If we assume further that 0 and ∞ are not eigenvalues of (A, B) , then we obtain the improved and sharp lower bound

$$n_R(A, B) \geq |n_+(B) - n_-(A)| + |n_-(A) - n_-(B)|. \quad (14)$$

The bound is a consequence of the intermediate value theorem, counting the number of eigenfunctions taking negative values at $t = \infty, 0, -\infty$, and noting (by assumption) that no nonzero eigenfunction is 0 at these three points. We have $|s(B)| = |n_+(B) - n_-(B)| \leq$

$|n_+(B) - n_-(A)| + |n_-(A) - n_-(B)|$ by the triangular inequality, and the sharpness of the bound (14) can be verified by taking

$$\hat{A} = \begin{bmatrix} R & 0 & 0 & 0 \\ 0 & D_1 & 0 & 0 \\ 0 & 0 & D_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} T & 0 & 0 & 0 \\ 0 & I_{n_+(B)-p} & 0 & 0 \\ 0 & 0 & -I_{n_-(B)-p} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where R, T, D_1, D_2 are as above with $p = \min(n_+(A), n_-(A), n_+(B), n_-(B))$. $\hat{A} - z\hat{B}$ has $n_+(B) + n_-(B) - 2p = r - 2p$ real eigenvalues. A direct calculation shows that $r - 2p = |n_+(B) - n_-(A)| + |n_-(A) - n_-(B)|$, as required.

If assuming that $0, \infty$ are not eigenvalues is undesirable, one can work with a shifted pencil: one analogously obtains $n_R(A, B) \geq |n_+(B - \epsilon A) - n_-(A - \epsilon B)| + |n_-(A - \epsilon B) - n_-(B - \epsilon A)|$ for an arbitrary $\epsilon > 0$, which is sharp assuming $\epsilon, 1/\epsilon$ are not eigenvalues and $(-1/\epsilon, 1/\epsilon)$ contains all finite real eigenvalues of (A, B) .

Remark 10 Some of the upper bounds in Theorem 5 are also sharp in the sense of Theorem 8. Specifically:

- The upper bounds in (7) and (9) are not always attainable. For example, if $n = 2m$ is even, $(n_+(A), n_0(A), n_-(A)) = (m, m, 0)$ and $(n_+(B), n_0(B), n_-(B)) = (0, m, m)$, our upper bound for $n_+(A, B)$ is n , but it is possible to prove with other means that no more than m positive eigenvalues, and indeed real eigenvalues can possibly be present. To see this, observe that in this case, by the same arguments as in the proof of Theorem 5 we have $n_0(A, B), n_\infty(A, B) \geq r - m$. Moreover, clearly $r \geq m$. Hence, $n_+(A, B) + n_-(A, B) \leq r - n_0(A, B) - n_\infty(A, B) \leq 2m - r \leq m$ (so trivially $n_+(A, B) \leq m$). Note that, generally, this technique of subtracting from r the lower bounds for the number of zero and infinite eigenvalues can be applied to compute alternative upper bounds for $n_+(A, B)$ or $n_-(A, B)$. Specifically, in the general case the above argument becomes $n_0(A, B) \geq n_0(A) - (n - r)$, $n_\infty(A, B) \geq n_0(B) - (n - r)$. So $n_+(A, B) + n_-(A, B) \leq r - n_0(A, B) - n_\infty(A, B) \leq r + 2(n - r) - n_0(A) - n_0(B) = 2n - r - n_0(A) - n_0(B)$. Such bounds can sometimes be stronger than the ones given in Theorem 5, but requires the rank r .
- The upper bounds in (8) and (11) are not always attainable. For example, suppose that $n = 2m$ is even and $(n_+(A), n_0(A), n_-(A)) = (n_+(B), n_0(B), n_-(B)) = (m, 0, m)$. Then, neither zero nor infinite eigenvalues can be present, but our upper bounds are equal to n .
- The upper bound in (10) is attained by the very same example as in the lower bound of (12).
- The upper bound in (12) is attained, for example, by any choice of \hat{A} and \hat{B} with both matrices diagonal and with $1, -1, 0$ elements on the main diagonal.

Although we opted in Theorem 5 not to give bounds that depend on the normal rank r , the latter can sometimes be computed (for example as $r = \max_i \text{rank}(A - \mu_i B)$ where μ_0, \dots, μ_n are any $n+1$ distinct complex numbers; in practice, with one randomly chosen $\mu_0 \in \mathbb{C}$, we have $r = \text{rank}(A - \mu_0 B)$ with high probability). In this scenario, the following

bounds, obtained from the intermediate steps in the proof of Theorem 5, are potentially more informative:

$$n_+(A, B) \leq 3n - r - n_0(A) - n_0(B) - N_{+-}, \quad (15)$$

$$n_0(A, B) \geq n_0(A) - n + r, \quad (16)$$

$$n_-(A, B) \leq 3n - r - n_0(A) - n_0(B) - N_{++}, \quad (17)$$

$$n_\infty(A, B) \geq n_0(B) - n + r. \quad (18)$$

In particular, if $r < n$ then (15) and (17) are always more informative, because for a singular pencil it holds

$$n - r < 2 \min(n_0(A), n_0(B)) = n_0(A) + n_0(B) - |\delta|.$$

(Proof of the last claim: $2 \max(\text{rank}(A), \text{rank}(B)) \leq 2r < n + r$, where the last inequality holds as we assume $r < n$.)

We conclude by pointing out a potentially practically important consequence of Theorem 5: if B is nearly positive definite, then the inertia of $A - zB$ is not very different from that of A . This is explored further in the next section.

4. Applications and extensions

4.1. Bounding the number of real eigenvalues in an interval

The classical Sylvester's law of inertia is useful for counting the number of real eigenvalues on an interval, say, (a, b) with $a < b$. For example, when $B \succ 0$, $|n_+(A - aB) - n_+(A - bB)|$ gives the exact number of the eigenvalues of (A, B) that lie in the interval (a, b) . We denote this number by $n_{(a,b)}(A, B)$. A similar notation is used for other types of subsets of the real line, e.g., $n_{[a,b]}(A, B)$ for the number of eigenvalues in the closed interval $[a, b]$. We also denote by $n_a(A, B)$ and $n_b(A, B)$ the number of eigenvalues of (A, B) that are equal to a and b respectively.

Here we investigate a possible extension of such results to a Hermitian indefinite pair (A, B) . Mathematically, this is a simple matter of applying a Möbius transformation [23, 29]: given $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ satisfying $\alpha\delta > \beta\gamma$, $\delta a = \beta$, $\gamma b = \alpha$ (it is always possible to find four such numbers provided that $a \neq b$), consider the map

$$A - zB \mapsto C - yD, \quad C := \delta A - \beta B, \quad D := \alpha B - \gamma A. \quad (19)$$

so that an eigenvalue of $A - zB$ in (a, b) is mapped to an eigenvalue of $C - yD$ in $(0, \infty)$. $n_+(C, D)$ is precisely equal to the number of eigenvalues of the pencil $A - zB$ in the interval (a, b) . We thus obtain

Corollary 11 *Let $a, b \in \mathbb{R}$ and A, B be Hermitian matrices, and denote the inertia of the matrix $A - aB$ by $(n_+(A - aB), n_0(A - aB), n_-(A - aB))$, and that of $bB - A$ by $(n_+(bB - A), n_0(bB - A), n_-(bB - A))$. Also define*

$$n_{++} = n_+(A - aB) + n_+(bB - A), \quad n_{--} = n_-(A - aB) + n_-(bB - A),$$

$$n_{+-} = n_+(A - aB) + n_-(bB - A), \quad n_{-+} = n_-(A - aB) + n_+(bB - A),$$

$$\delta = n_0(A - aB) - n_0(bB - A),$$

and $N_{++} = \max(n_{++}, n_{--})$, $N_{+-} = \max(n_{+-}, n_{-+})$. Then

$$\begin{aligned} N_{++} - n &\leq n_{(a,b)}(A, B) \leq 2n - |\delta| - N_{+-}, \\ \delta &\leq n_a(A, B) \leq 3n - N_{++} - N_{+-} - n_0(bB - A), \\ N_{+-} - n &\leq n_{(-\infty, a) \cup (b, \infty) \cup \infty}(A, B) \leq 2n - |\delta| - N_{++}, \\ 0 &\leq n_C(A, B) \leq 2 \min(n_+(A - aB), n_-(A - aB), n_+(bB - A), n_-(bB - A)), \\ -\delta &\leq n_b(A, B) \leq 3n - N_{++} - N_{+-} - n_0(A - aB), \\ |s(B)| &\leq n_R(A, B) \leq n - n_0(bB - A). \end{aligned}$$

In particular, we have the simple lower bound

$$\max(|n_+(A - aB) - n_+(A - bB)|, |n_-(A - aB) - n_-(A - bB)|) \leq n_{[a,b]}(A, B). \quad (20)$$

The proof is essentially a substitution $A \mapsto A - aB$, $B \mapsto bB - A$ in that of Theorem 8, and the bound (20) is a straightforward consequence of the intermediate value theorem. Note the interval is closed here, which is necessary.

Note that in this case the congruence transformation (in the sense discussed before Remark 7) applies to the matrices C and D , which are linear combinations of A, B .

4.2. Inertia of “nearly definite” matrix pencils

A typical case where we believe Theorem 5 would be useful is when B is “almost” definite, in that either most of its eigenvalues are positive or most are negative. Then the inertia of A gives a good estimate for that of (A, B) :

Corollary 12 *Let A, B be $n \times n$ Hermitian matrices, and suppose that B has $n - k$ positive eigenvalues. Then*

$$n - 2k \leq n_R(A, B) \leq n. \quad (21)$$

Furthermore, if k is small enough so that $n - 2k \geq |n_-(A) - n_+(A)|$, then

$$\begin{aligned} n_+(A) - k &\leq n_+(A, B) \leq n_+(A) + k, \\ n_-(A) - k &\leq n_-(A, B) \leq n_-(A) + k. \end{aligned} \quad (22)$$

Proof. The bounds follow from (7), (9) by direct calculations, taking $n_{++} \geq n_{--}$ and $n_{+-} \leq n_{-+}$ for (22), which hold by assumption. \square

Very similar results hold when either A or B is “almost” definite. We omit the details as the statements and proofs are essentially the same – recall that $A - zB$, $zB - A$, $B - zA$, and $zA - B$ all have the same number of positive eigenvalues and the same number of negative eigenvalues.

4.2.1. Parity of algebraic multiplicities

Assuming moreover that a, b are not eigenvalues of $A - \lambda B$, then for the algebraic multiplicities, we can further identify whether $n_{(a,b)}(A, B)$ can take even or odd values, as follows. For concreteness we treat the case where B has at least as many positive eigenvalues as negative ones; analogous results hold in the opposite case, and they can be derived by taking $B \mapsto -B$.

Theorem 13 *Let A, B be $n \times n$ Hermitian matrices, and suppose that B has $n - k$ positive eigenvalues, with $n \geq 2k$, and that neither a nor b is an eigenvalue of the pencil $A - zB$. Then,*

$$n_{(a,b)}(A, B) = |n_+(A - aB) - n_+(A - bB)| + 2h \quad (23)$$

for some $h \in \{0, 1, 2, \dots, k\}$. In particular,

$$|n_+(B) - n_-(B)| \leq n_R(A, B) \leq n. \quad (24)$$

Proof. It is clear that (24) can be obtained from (23) by taking $a = -b$ and letting $b \rightarrow +\infty$. It remains to prove (23).

Denote $a_+ = n_+(A - aB)$ and $b_+ = n_+(A - bB)$, and suppose that the rank of $A - zB$ is r . Then, $n - r$ eigenfunctions are identically zero. Recall [25, Sec. 5.2] that for a nonzero continuous function $f(x)$ such that $f(a)f(b) \neq 0$, the *local type of $f(x)$ in the interval $[a, b]$* is the ordered pair $(\text{sign } f(a), \text{sign } f(b))$. Denote by ℓ_τ the number of eigenfunctions of local type τ on $[a, b]$. Then we have the constraints

$$\ell_{(+,+)} + \ell_{(+,-)} = a_+, \quad \ell_{(+,+)} + \ell_{(-,+)} = b_+,$$

yielding

$$\ell_{(-,+)} - \ell_{(+,-)} = b_+ - a_+.$$

Noting that a root of even multiplicity corresponds to an even number of eigenvalues (counting algebraic multiplicities), [25, Prop. 5.5] implies that an eigenfunction corresponds to an odd number of eigenvalues (counting algebraic multiplicities) in (a, b) if and only if its local type is either $(+, -)$ or $(-, +)$. Hence, $n_{(a,b)}(A, B) - |a_+ - b_+|$ must be even, i.e., equal to $2h$ for some nonnegative integer h .

Suppose $h > k$ and let $t_0 > 0$ be large enough that the inertia of $A + t_0B$ and $A - t_0B$ are, respectively, $(n - k, n - r, r + k - n)$ and $(r + k - n, n - r, n - k)$. Using (20) with $a \mapsto -t_0, b \mapsto a$ and $a \mapsto b, b \mapsto t_0$, and noting that a, b are not eigenvalues, we see that the number of real eigenvalues is bounded below by

$$|n - k - a_+| + |a_+ - b_+| + |b_+ - k| + 2h \geq n - 2k + 2h > n,$$

a contradiction. Finally, to show $h \geq 0$, we can invoke the intermediate value theorem as before.

□

In particular, when k is known to be small, (23) shows that $|n_+(A - aB) - n_+(A - bB)|$ provides a good lower bound of $n_{(a,b)}(A, B)$. A summary of this subsection is that the inertia of Hermitian matrices $A - tB$ for $t \in \mathbb{R}$ gives some information about the real eigenvalues of (A, B) , but not sufficient to count the exact number of eigenvalues on an interval.

5. Nonlinear eigenvalue problems

We now turn to nonlinear eigenvalue problems. In this setting, we study $F(z)$, an analytic³ matrix-valued function of z , which takes Hermitian values for $z \in \mathbb{R}$. Again,

³In fact, our results are based on the intermediate value theorem, which holds more generally for continuous matrix-valued functions. We assume analyticity because (1) working over continuous functions requires several extra technicalities (2) in practice, functions of interest are analytic at least in appropriate regions of the complex plane.

its finite eigenvalues are defined as the values of $\lambda \in \mathbb{C}$ such that

$$\text{rank } F(\lambda) < \max_{z \in \mathbb{C}} \text{rank } F(z) =: r. \quad (25)$$

Examples include polynomial eigenvalue problems $F(z) = \sum_{i=0}^k z^i A_i$, $A_i \in \mathbb{C}^{n \times n}$. We assume $A_i = A_i^*$ are all symmetric (or Hermitian); such F are called symmetric (Hermitian) matrix polynomials. As in the linear case, we formally define $F(\infty) := A_k$ and seek eigenvalues in \mathbb{C}^* . As before, we explore Sylvester-like inertia laws for F , or rather, lower bounds for the number of eigenvalues in an interval (a, b) . Obtaining upper bounds appears to be not possible, at least by the approach presented here: indeed, the total number of eigenvalues of $F(z)$ cannot be bounded in general, and in fact it might be not even countable if $F(z)$ is not analytic. There are exceptions, e.g., when $F(z)$ is polynomial (in which case the number of eigenvalues is at most its rank times its degree [8]). It is sometimes possible to quantify the number of eigenvalues in a given domain when $F(z)$ is assumed analytic [4] or, again, polynomial [5, 30].

Along the lines of the previous sections, we use the following straightforward result.

Proposition 14 *Let $F(z)$ be a Hermitian matrix function, with elements that depend analytically in z . Let $a < b \in \mathbb{R}$ and suppose that neither a nor b is an eigenvalue of $F(z)$. Then, the nonlinear eigenvalue problem $F(z)$ has at least $|n_+(F(a)) - n_+(F(b))|$ real eigenvalues in the open interval (a, b) .*

Proof. Recalling Theorem 2, we examine the n eigenfunctions $\lambda_i(t)$, which are continuous: $n_+(F(a))$ of them must be positive at $t = a$ and $n_+(F(b))$ of them are positive at $t = b$. Hence, at least $|n_+(F(a)) - n_+(F(b))|$ of them must cross the axis of the abscissae at least once. \square

Proposition 14, while elementary to prove, can be used to show easily that the class of *definite matrix polynomials* introduced in [12] have only real eigenvalues, and also to determine their types, which are defined so that they correspond to the signs of the lowest order nonzero derivatives of $\lambda_i(t)$ at the zeros. Such matrix polynomials ($F(z) := P(z) = \sum_{i=0}^k z^i A_i$ of degree k are those for which there exist $\{\mu_i\}_{i=1}^{k+1}$ such that $(-1)^i P(\mu_i)$ is positive (or negative) definite for all i . Then, applying Theorem 14 with $a \mapsto \mu_i, b \mapsto \mu_{i+1}$ shows P has (at least) n eigenvalues in (μ_i, μ_{i+1}) . Since this holds for all i , it follows that all the nk eigenvalues are real; and hence each interval (μ_i, μ_{i+1}) contains exactly n eigenvalues. We can further derive the sign characteristics of the eigenvalues as the sign of the derivatives of $\lambda_i(t)$ at the zeros, which are all the same for the n eigenvalues in each interval (and the signs must differ between neighboring intervals).

Proposition 14 has little resemblance to the original Sylvester's law of inertia, as there is no congruence transformations involved. A weak result can be obtained along this line when P is a matrix polynomial $P(z) = \sum_{i=0}^k z^i A_i$. Specifically, take $a = 0, b = \infty$. Then $|n_+(P(a)) - n_+(P(b))|$ is determined solely by the leading and trailing coefficients A_0 and A_k , since $n_+(P(a)) = n_+(A_0)$ and $n_+(P(b)) = n_+(A_k)$. Since $n_+(A_0), n_+(A_k)$ in turn are invariant under congruence transformations $A_0 \mapsto X^* A_0 X, A_k \mapsto Y^* A_k Y$, it follows that any matrix polynomial of the form $\tilde{P}(z) = z^k Y^* A_k Y + X^* A_0 X + \sum_{i=1}^{k-1} z^i A_i$ has at least $|n_+(A_0) - n_+(A_k)|$ real positive eigenvalues where A_i for $i = 1, \dots, k-1$ are arbitrary Hermitian matrices and X, Y are arbitrary invertible matrices. Similarly, $\tilde{P}(z)$

has at least $|n_+(A_0) - n_+(A_k)|$ (resp. $|n_+(A_0) - n_-(A_k)|$) real negative eigenvalues if k is even (resp. odd).

6. Geometric multiplicities

Our results so far have been concerned with the algebraic multiplicities of eigenvalues. We now comment on the geometric counterparts. We recall (see, e.g., [25, Def. 2.3]) that for an analytic matrix function the partial multiplicities of a finite eigenvalue λ are the multiplicities of λ as a root of the nonzero Rellich eigenfunctions; the algebraic multiplicity of λ is the sum of the partial multiplicities while the geometric multiplicity of λ is the number of its nonzero partial multiplicities. It transpires that most bounds are valid for the geometric multiplicities: most of the proofs made no reference to the fact that the algebraic multiplicity is counted. Therefore, bounds in the following are valid also for geometric multiplicities: Theorem 5, the bounds (15)–(18), Corollaries 11 and 12, and Proposition 14.

An exception is Theorem 13. There, the argument in connection with the multiplicities of the roots of the eigenfunctions apply specifically to the algebraic multiplicities. For the geometric multiplicity, a “less informative” bound holds:

$$|n_+(A - aB) - n_+(A - bB)| \leq n_{(a,b)}(A, B) \leq |n_+(A - aB) - n_+(A - bB)| + 2k. \quad (26)$$

The lower bound is essentially (20), and the upper bound is a direct consequence of (23) and the fact that the algebraic multiplicity is an upper bound for the geometric multiplicity.

It is worth noting that multiple eigenvalues are a nongeneric phenomenon, which do not arise e.g. for random symmetric matrices A, B . In such cases the two multiplicities are clearly the same, and the formula (23) is valid for both types of multiplicities.

7. Experiments

7.1. Number of real eigenvalues for B nearly definite

We first examine generalized indefinite eigenvalue problems $Ax = \lambda Bx$ where B is nearly definite.

In Figure 3 we plot the Rellich’s eigenfunctions $\lambda_i(t)$ for $F(t) = A - tB$ with $A = A^T, B = B^T \in \mathbb{R}^{7 \times 7}$ and B indefinite. In both figures, the inertia of B is $(6, 0, 1)$. We verify the bounds in (12) hold, which tell us that (A, B) has at least 5 and at most 7 real eigenvalues.

Note also the evidently observable phenomenon of eigenvalue repulsion effect in Figures 3, which has been the hallmark of eigenvalue theory in the field of symmetric parameterized eigenvalue problems [19, Sec. 9.5]. We emphasize the fact that this does not hold if the parameterized matrices are nonsymmetric.

7.2. Variation of the inertia of (A, B) when the inertia of A, B is fixed

Next we set B as before, but now fix also the inertia A to $(5, 0, 2)$. By Theorem 1, all such A ’s are congruent to each other, and so are all such B ’s. As before, such (A, B) must have at least five real eigenvalues, but more can be said: Theorem 5 shows that (A, B)

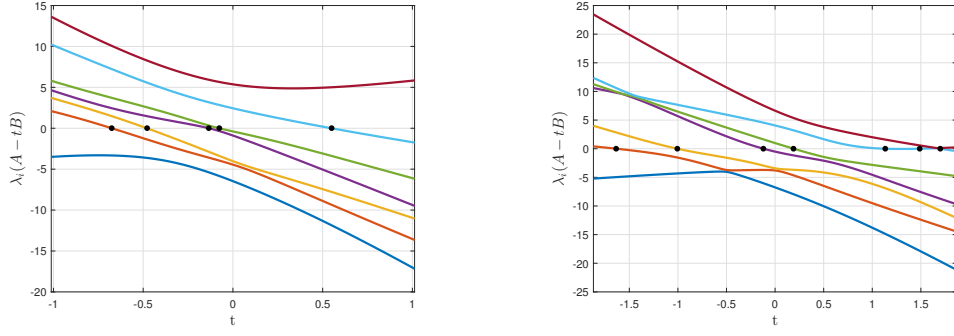


Figure 3: $\lambda_i(t)$ for two pencils $A - tB$ with inertia of B fixed to $(n_+(B), n_0(B), n_-(B)) = (6, 0, 1)$. The left pencil has five real eigenvalues, while the right has $n = 7$.

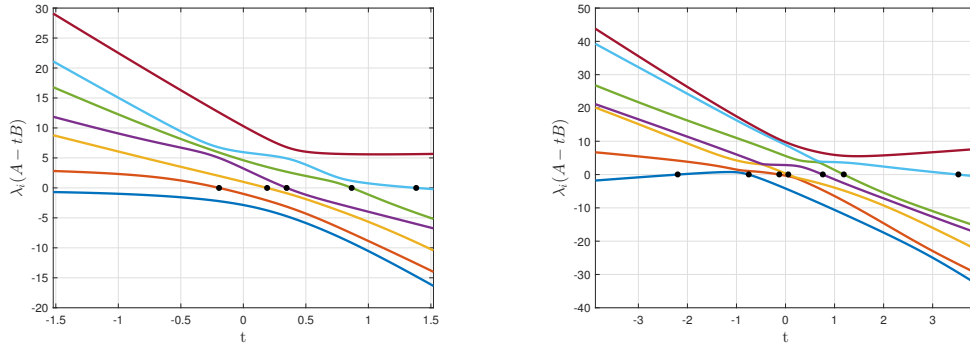


Figure 4: Same setting as in Figure 3, but with inertia of A fixed to $(n_+(A), n_0(A), n_-(A)) = (5, 0, 2)$. The left pencil has four positive and one negative eigenvalues, while all the eigenvalues of the right pencil are real.

has at least four positive eigenvalues, and one negative eigenvalue. Figure 4 illustrates this and the sharpness of the bounds.

Using the bound (23), we can obtain significant information on the number of eigenvalues lying in intervals. For example, by computing the inertia of $A - tB$ at $t = -\frac{1}{2}$, which is $(6, 0, 1)$, together with that of B , we see that both pencils in Figure 4 satisfy $n_{(t, \infty)} \in \{5, 7\}$ and $n_{(-\infty, t)} \in \{0, 2\}$ (the algebraic multiplicity). The simple bound (20) is sharp in the left pencil, and $h = 1$ in (23) for $n_{(t, \infty)}$ of the right pencil.

7.3. Estimating the number of eigenvalues in an interval

Equation (23) in Theorem 13 gives the possible number of eigenvalues of $A - \lambda B$ that lie in an interval (a, b) from $n_+(A - aB)$, $n_+(A - bB)$ and $n_+(B)$ (we make the generic assumption that the eigenvalues are simple). Here we illustrate the result and examine a typical value of $2h$ in (23). We generated random symmetric matrices A, B by the MATLAB⁴ code `X = randn(n); X = X+X'`, and then shift $B \mapsto B - sI$, where (denoting by $\mu_i(B)$ the i th eigenvalue of the matrix B , ordered in a non-decreasing manner) $s = (\mu_k(B) + \mu_{k+1}(B))/2$, so that B has exactly k negative eigenvalues. We then compute the bound (23), along with the exact number s of eigenvalues in (a, b) , and compute $(2h :=) s - |n_+(A - aB) - n_+(A - bB)|$, the underestimation count of the lower bound $|n_+(A - aB) - n_+(A - bB)|$. We took $n = 5000$, $(a, b) = (0, 1)$ and generated 1000 such random instances, and Figure 5 shows their histograms for $k = 10$ (left) and $k = 100$ (right). We observe that a typical value of h lies around $0.3k$ in this configuration, and the tail of the distribution becomes lighter as k grows. As long as n is large enough, the figures are largely independent of the matrix size n ; for example, $n = 1000$ gave very similar results.

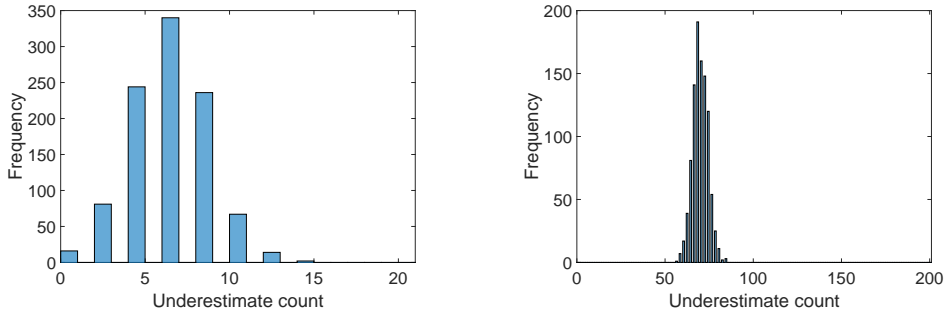


Figure 5: Histograms from 1000 randomly generated instances of $2h \in \{0, 2, 4, \dots, 2k\}$ (left: $k = 10$, right: $k = 100$) in Theorem 13, the difference between the exact number of eigenvalues and its lower bound $|n_+(A - aB) - n_+(A - bB)|$. $A, B \in \mathbb{R}^{5000 \times 5000}$.

7.4. Quadratic matrix polynomials

Hyperbolic case. We now turn to hyperbolic quadratic matrix polynomials $P(z) = z^2A + zB + C$ with $A, B, C \in \mathbb{R}^{n \times n}$ symmetric and $0 \succ A, C \succ 0$, which have only real

⁴MATLAB is a registered trademark by The MathWorks, Inc.

eigenvalues (a special case of definite matrix polynomials). As discussed in Section 5, it is easy to see why the eigenvalues are all real: the n Rellich's eigenfunctions $\lambda_i(t)$ (setting $F(t) = At^2 + Bt + C$ in Theorem 2) are all positive for $t = 0$ and all negative for $t = \pm M$, for any large enough positive constant M . Figure 6 plots the eigenfunctions for an example with $n = 5$.

In fact, for any interval (a, b) , assuming that a, b are not eigenvalues, one can compute the exact number of eigenvalues in it from the inertia of $P(a)$ and $P(b)$: for example when $a < 0 < b$, $n_-(P(b)) + n_-(P(a))$ is the number of eigenvalues. This is because precisely $n_-(P(a))$ eigenfunctions still have to cross the horizontal axis between a and 0, and $n_-(P(b))$ must have crossed the horizontal axis between 0 and b . In particular, we can bound the largest (or smallest) eigenvalue $\lambda_{\max}(P)$ (or $\lambda_{\min}(P)$) of P as follows: examine the inertia of $P(t) = t^2A + tB + C$ for $t > 0$ ($t < 0$), and if they are all negative, then $\lambda_{\max}(P) < t$ ($\lambda_{\min}(P) > t$).

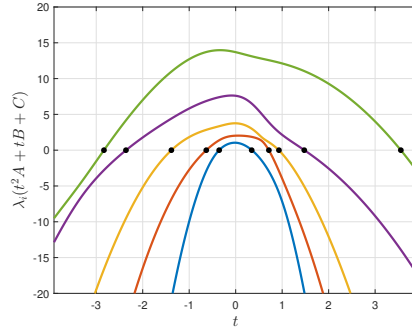


Figure 6: Rellich's eigenfunctions for a 5×5 quadratic hyperbolic $P(t) = At^2 + Bt + C$. Note that $\lambda_i(t)$ are all negative for sufficiently small t , all positive at $t = 0$, and all negative again at sufficiently large t .

NLEVP's spring example. We consider the quadratic polynomial eigenvalue problem $P(\lambda) = \lambda^2A + \lambda B + C$ in NLEVP [2] `nlevp('spring', n, 1, 10*ones(n, 1))`, with matrices

$$A = I, \quad B = \beta \begin{bmatrix} 20 & -10 & & & \\ -10 & 30 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 30 & -10 \\ & & & -10 & 20 \end{bmatrix} \quad C = \begin{bmatrix} 15 & -10 & & & \\ -10 & 15 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 15 & -10 \\ & & & -10 & 15 \end{bmatrix}.$$

As explained in [11], for values of β larger than ≈ 0.52 , P is definite, but it is indefinite for smaller values of β . For example, taking $\beta = 0.3$ and $n = 7$, we plot the eigenvalues of $F(t) = At^2 + Bt + C$ in Figure 7. The inertiae of the matrix $F(t)$ at $t = -13$ and $t = -4$ are $(7, 0, 0)$ and $(3, 0, 4)$, respectively. Thus Proposition 14 guarantees that P has at least 4 eigenvalues in the interval $(-13, -4)$. This bound happens to be sharp here; the effectiveness of this approach evidently depends on the choice of the interval (for example, taking the inertia at $t = -15$ and $t = 0$ we obtain no useful information).

The same approach can be used in larger scale. For example, taking $n = 1000$ in the same example, and examining the inertia of $F(t)$ at $t = -14$ and -3.3 , we see that

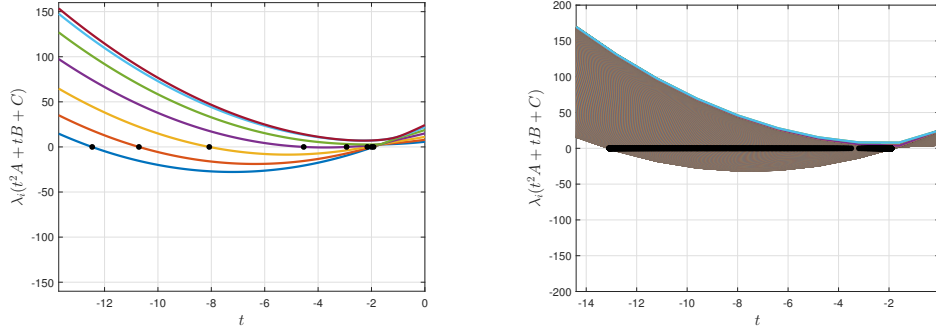


Figure 7: Rellich's eigenfunctions for NLEVP's **spring** problem. $F(t) = At^2 + \beta Bt + C$, $\beta = 0.3$, $n = 7$ (left) and $n = 1000$ (right). This is an indefinite example, but by Proposition 14 we can obtain nontrivial information on the number of eigenvalues in an interval.

there are at least 626 eigenvalues in the interval $(-14, -3.3)$, which also happens to be the correct value (see Figure 7 (right)). With the current fastest algorithms available for each task, computing the inertia of $F(t)$ at two values of t (which can be done via computing two LDL^T factorization of $n \times n$ symmetric matrices) is more efficient than computing the eigenvalues of P (for which a standard method requires the solution of a $2n \times 2n$ eigenvalue problem). One can also obtain estimates of the inertia of Hermitian matrices more efficiently by estimating the spectral density [20].

7.5. When Jordan blocks are present

In all experiments so far, the eigenvalues are simple, and hence the algebraic and geometric multiplicities are always the same. Here we examine the case when they are different, that is, when Jordan blocks of size two or larger are present. For example, we consider a pencil $A - tB$ in Hermitian canonical form and having a single Jordan block associated with the real finite eigenvalue λ :

$$A = \begin{bmatrix} & & 1 & \lambda \\ & & & \lambda \\ & \ddots & \ddots & \\ 1 & \lambda & & \\ \lambda & & & \end{bmatrix}, \quad B = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \end{bmatrix}. \quad (27)$$

In such cases, the eigenfunctions have zeros of high multiplicity, and Theorem 5 is not tight. We illustrate this in Figure 8, where A, B are the 6×6 case of (27) for $\lambda = 1$. Here the plots of the eigenfunctions of $A - tB$ reveal the eigenvalue 1 and its geometric multiplicity 1, but it is impossible from the figure alone to identify the algebraic multiplicity 6 (for this goal, we would need to study the plots of the first, second, \dots , fifth derivatives of the $\lambda_i(t)$ as well). Theorem 5 gives only trivial bounds $0 \leq n_+(A, B) \leq n$, $0 \leq n_-(A, B) \leq n$. The inertia of $A - tB$ is $(3, 0, 3)$ for any value of $t \neq 1$, and our results, e.g., (26) provide no information on the geometric multiplicities of the eigenvalues from the inertia of $A - tB$, unless we take $t = 1$. Theorem 13 shows that the algebraic multiplicities of real eigenvalues (if any) are even. Recall that (23) in Theorem 13 is not satisfied for the geometric multiplicity.

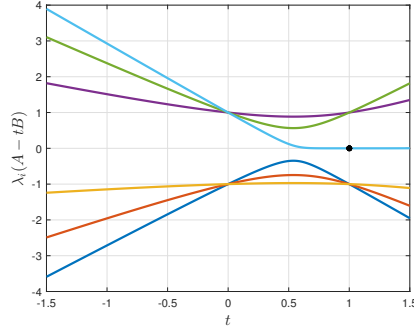


Figure 8: Rellich's eigenfunctions of the $A - tB$ for A, B as in (27), a 6×6 pair with Jordan block of size 6 at $\lambda = 1$.

We note that, unlike the previous plots, the curves intersect at $t = 0$ and $t = 1$ in this example. This nongeneric behavior happens because A, B are in such special forms such that the pencil $A - zB$ has a Jordan block.

Acknowledgement

We thank Andy Wathen for pointing out the role played by eigenvalues of indefinite pencils in constraint preconditioners. We thank two anonymous referees for their constructive and detailed comments, that helped us to improve the presentation. The second author is supported by JSPS grants no. 17H01699 and 18H05837.

- [1] S. Adachi, S. Iwata, Y. Nakatsukasa, and A. Takeda. Solving the trust-region subproblem by a generalized eigenvalue problem. *SIAM J. Optim.*, 27(1):269–291, 2017.
- [2] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur. Nlevp: A collection of nonlinear eigenvalue problems. *ACM Trans. Math. Soft.*, 39(2):7, 2013.
- [3] B. Bilir and C. Chicone. A generalization of the inertia theorem for quadratic matrix polynomials. *Linear Algebra Appl.*, 280:229–240, 1998.
- [4] D. Bindel and A. Hood. Localization theorems for nonlinear eigenvalue problems. *SIAM Rev.*, 57(4):585–607, 2015.
- [5] D. A. Bini, V. Noferini, and M. Sharify. Locating the eigenvalues of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 34(4):1708–1727, 2013.
- [6] M. I. Bueno, F. M. Dopico, S. Furtado, and M. Rychnovsky. Large vector spaces of block-symmetric strong linearizations of matrix polynomials. *Linear Algebra Appl.*, 477:165–210, 2015.
- [7] L. De Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM J. Matrix Anal. Appl.*, 28(3):642–666 (electronic), 2006.
- [8] F. De Terán, F. M. Dopico, and D. S. Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. *Linear Algebra Appl.*, 459:264–333, 2014.
- [9] J. Demmel. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, USA, 1997.
- [10] E. Di Napoli, E. Polizzi, and Y. Saad. Efficient estimation of eigenvalue counts in an interval. *Numer. Lin. Alg. Appl.*, 23(4):674–692, 2016.
- [11] C.-H. Guo, N. J. Higham, and F. Tisseur. An improved arc algorithm for detecting definite Hermitian pairs. *SIAM J. Matrix Anal. Appl.*, 31(3):1131–1151, 2009.
- [12] N. J. Higham, D. S. Mackey, and F. Tisseur. Definite matrix polynomials and their linearization by definite pencils. *SIAM J. Matrix Anal. Appl.*, 31(2):478–502, 2009.
- [13] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, second edition, 2012.
- [14] K. D. Ikramov. On the inertia law for normal matrices. *Dokl. Math.*, 64:141142, 2001.
- [15] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, 2nd edition, 1966.
- [16] C. Keller, N. I. M. Gould, and A. J. Wathen. Constraint preconditioning for indefinite linear systems. *SIAM J. Matrix Anal. Appl.*, 21(4):1300–1317, 2000.

- [17] A. Kostić and H. Voss. On Sylvester's law of inertia for nonlinear eigenvalue problems. *Electron. Trans. Numer. Anal.*, 40:82–93, 2013.
- [18] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Rev.*, 47(3):407–443, 2005.
- [19] P. Lax. *Linear Algebra and Its Applications*. Wiley, 2nd edition, 2007.
- [20] L. Lin, Y. Saad, and C. Yang. Approximating spectral densities of large matrices. *SIAM Rev.*, 58(1):34–65, 2016.
- [21] W. W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302–303:469–533, 1999.
- [22] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28:971–1004, 2005.
- [23] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Möbius transformations of matrix polynomials. *Linear Algebra Appl.*, 470:120–184, 2015.
- [24] A. S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*. American Mathematical Soc., 2012.
- [25] V. Mehrmann, V. Noferini, F. Tisseur, and H. Xu. On the sign characteristics of Hermitian matrix polynomials. *Linear Algebra Appl.*, 511:328–364, 2016.
- [26] V. Mehrmann and H. Xu. Structured Jordan canonical forms for structured matrices that are Hermitian, skew Hermitian or unitary with respect to indefinite inner products. *Electron. J. Linear Algebra*, 5:67–104, 1999.
- [27] Y. Nakatsukasa, V. Noferini, and A. Townsend. Computing the common zeros of two bivariate functions via Bézout resultants. *Numer. Math.*, 129:181–209, 2015.
- [28] Y. Nakatsukasa, V. Noferini, and A. Townsend. Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach. *SIAM J. Matrix Anal. Appl.*, 38(1):1–29, 2017.
- [29] V. Noferini. The behavior of the complete eigenstructure of a polynomial matrix under a generic rational transformation. *Electron. J. Linear Algebra*, 23:607–624, 2012.
- [30] V. Noferini, M. Sharify, and F. Tisseur. Tropical roots as approximations to eigenvalues of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 36(1):138–157, 2015.
- [31] F. Rellich. Störungstheorie der Spektralzerlegung I. *Math. Anal.*, 113:600–619, 1937.
- [32] F. Rellich. *Perturbation Theory of Eigenvalue Problems*. Gordon and Breach, 1969.
- [33] J. J. Sylvester. A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares. *Philosoph. Mag.*, 4(32):138142, 1852.
- [34] R. C. Thompson. The characteristic polynomial of a principal subpencil of a hermitian matrix pencil. *Linear Algebra Appl.*, 14:135–177, 1976.