Electronic Transactions on Numerical Analysis.
Volume 60, pp. 276–291, 2024.

ISSN 1068–9613.
DOI: 10.1553/etna_vol60s276

ETNA
Kent State University and
Johann Radon Institute (RICAM)

SPECTRAL PROPERTIES OF CERTAIN NONSYMMETRIC SADDLE POINT MATRICES*

JÖRG LIESEN† AND JUSTUS RAMME†

Abstract. We consider certain (real) nonsymmetric matrices in saddle point form, study their general Jordan normal forms, and prove new conditions so that these matrices are diagonalizable with a real spectrum. For matrices satisfying our conditions we show how to construct an inner product in which these matrices are selfadjoint. Our approach generalizes previously published results in this area, which require stronger assumptions on the given saddle point matrices and hence are less widely applicable.

Key words. saddle point problems, eigenvalues and eigenvectors, conjugate gradient iterations, Krylov subspace methods

AMS subject classifications. 15A18, 65F10

1. Introduction.

Linear algebraic systems in saddle point form, i.e.,

$$\mathcal{S}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad \text{where } \mathcal{S} := \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}, \ A = A^T \in \mathbb{R}^{n,n}, \ C = C^T \in \mathbb{R}^{m,m}, \tag{1.1}$$

arise in numerous applications in science and engineering; see [1] for an extensive survey. Often it is assumed that A is positive definite ($A > 0$) and that C is positive semidefinite ($C \geq 0$). By decomposing

$$\mathcal{S} = \begin{bmatrix} I_n & 0 \\ BA^{-1} & I_m \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix} \begin{bmatrix} I_n & A^{-1}B^T \\ 0 & I_m \end{bmatrix},$$

where $S := C + BA^{-1}B^T$, it follows that $\mathcal{S}$ has n positive, $\operatorname{rank}(S)$ negative, and $m - \operatorname{rank}(S)$ zero eigenvalues. Unless m is very small or S is highly rank-deficient, which is both rare in practical applications, the saddle point matrix $\mathcal{S}$ is highly indefinite. This is an unfortunate circumstance, since many iterative methods, and in particular Krylov subspace methods, converge rather slowly for highly indefinite problems; see, e.g., [1, Section 9.2].

As one possible remedy it was suggested by several authors to multiply the second block row in (1.1) by -1 and hence to consider the equivalent linear algebraic system

$$\mathcal{A}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ -b_2 \end{bmatrix}, \quad \text{where } \mathcal{A} := \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix}, \ A = A^T \in \mathbb{R}^{n,n}, \ C = C^T \in \mathbb{R}^{m,m}; \tag{1.2}$$

see [1, pp. 23–27]. The nonsymmetric matrix $\mathcal{A}$ can be written as

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} + \begin{bmatrix} 0 & B^T \\ -B & 0 \end{bmatrix}, \tag{1.3}$$

which is the splitting of $\mathcal{A}$ into its symmetric and skew-symmetric parts. If $A > 0$ and $C \geq 0$, then the symmetric part is positive semidefinite, and it immediately follows that the eigenvalues of $\mathcal{A}$ have nonnegative real parts, i.e., $\operatorname{Re}(\sigma(\mathcal{A})) \geq 0$. Moreover, if $\mathcal{A}$ is nonsingular, then $\mathcal{A}$ is (nonsymmetric) positive definite. We point out that the formulation of the problem with the "natural" splitting into (1.3) with a positive semidefinite symmetric part is typical for

*Received April 3, 2024. Accepted April 4, 2024. Published online on May 14, 2024. Recommended by Lothar Reichel.

†Institute of Mathematics, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany ({liesen,ramme}@math.tu-berlin.de).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

discretized problems that arise from dissipative Hamiltonian DAE systems. This topic has attracted a lot of attention in recent years; see, e.g., [8] and the references given therein.

Several authors have studied spectral properties of $\mathcal{A}$. For example, a complete analysis of $\sigma(\mathcal{A})$ for the case $A = \eta I_n$ and $C = 0$ is given in [6], and bounds for the eigenvalues of $\mathcal{A}$ for the case $A > 0$ and $C > 0$ are given in [3]. Another line of work deals with conditions for $\mathcal{A}$ so that this matrix is diagonalizable with a positive real spectrum. If this holds, then a conjugate gradient method for $\mathcal{A}$ can be constructed. This method may converge faster than a Krylov subspace method applied to the system with the symmetric indefinite matrix $\mathcal{S}$.

In [2] the authors derived conditions for diagonalizability of $\mathcal{A}$ with a real spectrum for the case $A > 0$ and $C = 0$; see in particular [2, Proposition 3.1]. The results from [2] were extended to the case $A > 0$ and $C \geq 0$ in [10]. Both approaches are based on the observation that

$$\mathcal{M}(\gamma)\mathcal{A} = \mathcal{A}^T\mathcal{M}(\gamma), \quad \text{where} \quad \mathcal{M}(\gamma) := \begin{bmatrix} A - \gamma I_n & B^T \\ B & \gamma I_m - C \end{bmatrix}, \quad \gamma \in \mathbb{R}. \tag{1.4}$$

If the parameter γ can be chosen so that $\mathcal{M}(\gamma) > 0$, then $\mathcal{M}(\gamma)$ defines an inner product on $\mathbb{R}^{n+m}$ in which $\mathcal{A}$ is symmetric. This is a *sufficient* condition so that $\mathcal{A}$ is diagonalizable with a real spectrum. A *necessary* condition for $\mathcal{M}(\gamma) > 0$ is that $\lambda_{\min}(A) > \gamma > \lambda_{\max}(C)$; see [10, Theorem 2.2]. Consequently, any sufficient condition for diagonalizability of $\mathcal{A}$ with a real spectrum derived in this way requires that $\lambda_{\min}(A) > \lambda_{\max}(C)$. For a related line of research on Krylov subspace methods with nonstandard inner products for matrices in saddle point form, we refer to [4, 12, 15].

In this paper we study spectral properties of $\mathcal{A}$, starting with general results about the Jordan decomposition of $\mathcal{A}$ that are derived using the theory of indefinite linear algebra. In particular, we give bounds for the sizes of the largest possible Jordan blocks that can occur in the Jordan normal form of $\mathcal{A}$, and we give examples of matrices $\mathcal{A}$ for which these bounds are attained. We then present an alternative and direct approach to the spectral analysis of $\mathcal{A}$ that avoids the detour via the matrix $\mathcal{M}(\gamma)$. In this way we are able to derive conditions for diagonalizability of $\mathcal{A}$ with a real spectrum that generalize the previously published results. In particular, we are able to replace the restrictive condition $\lambda_{\min}(A) > \lambda_{\min}(C)$ by a condition for the minimum distance of $\sigma(A)$ and $\sigma(C)$. Thus, our approach allows that $\sigma(A)$ and $\sigma(C)$ interlace (though not intersect). We also explain how in this case an inner-product matrix for $\mathcal{A}$ may be constructed, generalizing the construction with the matrix $\mathcal{M}(\gamma)$ that was given in [2, 10].

2. General results about the Jordan decomposition of $\mathcal{A}$. In order to analyze the Jordan decomposition of $\mathcal{A}$ we will use results from the theory of indefinite linear algebra; see [7] for a comprehensive treatment.

Let $H \in \mathbb{C}^{n,n}$ be Hermitian and nonsingular. A matrix $M \in \mathbb{C}^{n,n}$ is called *H-selfadjoint* when

$$M = H^{-1}M^H H.$$

Note that $\mathcal{A}$ is $\mathcal{M}(\gamma)$-selfadjoint if $\mathcal{M}(\gamma)$ is nonsingular (see (1.4)), and if we define

$$\mathcal{J} := \begin{bmatrix} I_n & 0 \\ 0 & -I_m \end{bmatrix},$$

then $\mathcal{J} = \mathcal{J}^T = \mathcal{J}^{-1}$ and $\mathcal{J}\mathcal{A} = \mathcal{A}^T\mathcal{J}$, i.e., $\mathcal{A}$ is $\mathcal{J}$-selfadjoint.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

For $\lambda \in \mathbb{C}$ and $s \geq 1$ we denote by $J_s(\lambda)$ the Jordan block with eigenvalue λ of size s. If $\mathrm{Im}(\lambda) > 0$ and $s \geq 2$ is even, we define

$$\widetilde{J}_s(\lambda) := \begin{bmatrix} J_{\frac{s}{2}}(\lambda) & 0 \\ 0 & J_{\frac{s}{2}}(\overline{\lambda}) \end{bmatrix}.$$

With this notation we can state the following important result; see [7, Theorem 5.1.1].

THEOREM 2.1. *If $M \in \mathbb{C}^{n,n}$ is H-selfadjoint, then there exists a nonsingular $X \in \mathbb{C}^{n,n}$ such that $M = X^{-1}JX$ and $H = X^H P X$, where*

$$J = \mathrm{diag}\left(J_{n_1}(\lambda_1) \ldots J_{n_\alpha}(\lambda_\alpha), \widetilde{J}_{n_{\alpha+1}}(\lambda_{\alpha+1}), \ldots, \widetilde{J}_{n_\beta}(\lambda_\beta)\right) \tag{2.1}$$

is a Jordan normal form of A with

$$\lambda_j \in \mathbb{R}, \ \textit{for } j = 1, \ldots, \alpha, \quad \textit{and} \quad \mathrm{Im}(\lambda_j) > 0, \ \textit{for } j = \alpha+1, \ldots, \beta.$$

Moreover,

$$P = \mathrm{diag}\left(\varepsilon_1 I^r_{n_1}, \ldots, \varepsilon_\alpha I^r_{n_\alpha}, I^r_{n_{\alpha+1}}, \ldots, I^r_{n_\beta}\right), \ \textit{where } I^r_k = \begin{bmatrix} 0 & \ldots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \ldots & 0 \end{bmatrix} \in \mathbb{R}^{k,k},$$

and $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_\alpha\}$ is an ordered set of signs ± 1 that is uniquely determined by M and H up to permutation of signs corresponding to equal Jordan blocks.

The Jordan normal form (2.1) shows, in particular, that the non-real eigenvalues of an H-selfadjoint matrix M appear in complex conjugate pairs and that the sizes of the Jordan blocks with eigenvalue λ are equal to the sizes of the Jordan blocks with eigenvalue $\overline{\lambda}$.

The next result is originally from [2, Proposition 2.3] and gives an upper bound for the maximal number of non-real eigenvalues of the matrix $\mathcal{A}$. We give an alternative proof based on Theorem 2.1.

THEOREM 2.2. *A matrix $\mathcal{A}$ as in (1.2) can have at most $2m$ non-real eigenvalues, counting conjugates.*

Proof. Since $\mathcal{A}$ is $\mathcal{J}$-selfadjoint we have $\mathcal{A} = X^{-1}JX$ and $\mathcal{J} = X^H P X$ for some nonsingular matrix X by Theorem 2.1. The matrices $\mathcal{J}$ and P are congruent, which shows that P has exactly m negative and n positive eigenvalues. The non-real eigenvalues of $\mathcal{A}$ are contained in the diagonal blocks $J_{n_j}(\lambda_j)$, $j = \alpha+1, \ldots, \beta$, in J, where $n_j \geq 2$. Each complex conjugate pair of non-real eigenvalues of $\mathcal{A}$ thus corresponds to a diagonal block $I^r_{n_j} \in \mathbb{R}^{n_j, n_j}$, $j = \alpha+1, \ldots, \beta$, in P. Each such matrix $I^r_{n_j}$ has at least one negative eigenvalue (namely -1), which shows that there can be at most m such matrices and thus at most $2m$ non-real eigenvalues of $\mathcal{A}$, counting conjugates. ∎

In the following result we characterize the largest Jordan block sizes that can occur in the Jordan normal form of $\mathcal{A}$.

THEOREM 2.3. *If $J_s(\lambda)$ is a Jordan block of a matrix $\mathcal{A}$ as in (1.2), then $s \leq 2m+1$ if $\lambda \in \mathbb{R}$, and $s \leq m$ if $\lambda \in \mathbb{C} \setminus \mathbb{R}$.*

Proof. By Theorem 2.1 we have $\mathcal{A} = X^{-1}JX$ and $\mathcal{J} = X^H P X$, where

$$J = \mathrm{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_\alpha}(\lambda_\alpha), \widetilde{J}_{n_{\alpha+1}}(\lambda_{\alpha+1}), \ldots, \widetilde{J}_{n_\beta}(\lambda_\beta)), \tag{2.2}$$

$$P = \mathrm{diag}(\varepsilon_1 I^r_{n_1}, \ldots, \varepsilon_\alpha I^r_{n_\alpha}, I^r_{n_{\alpha+1}}, \ldots, I^r_{n_\beta}). \tag{2.3}$$

The matrices $J_{n_j}(\lambda_j)$, $j = 1, \ldots, \alpha$, are the Jordan blocks of $\mathcal{A}$ corresponding to the real eigenvalues. Each of them corresponds to a diagonal block $\varepsilon_j I^r_{n_j}$ of P. The matrix $I^r_{n_j}$ has

ETNA
Kent State University and
Johann Radon Institute (RICAM)

exactly $\lfloor \frac{n_j}{2} \rfloor$ eigenvalues -1, and $\lceil \frac{n_j}{2} \rceil$ eigenvalues $+1$. If $n_j \geq 2m+2$, then $\lfloor \frac{n_j}{2} \rfloor \geq m+1$ and $\lceil \frac{n_j}{2} \rceil \geq m+1$, which means that regardless of the sign of ε_j, the matrix $\varepsilon_j I_{n_j}^r$ will have at least $m+1$ negative eigenvalues. But this is a contradiction since P and $\mathcal{J}$ are congruent and $\mathcal{J}$ has exactly m negative eigenvalues. Consequently, $n_j \leq 2m+1$.

The matrices $\widetilde{J}_{n_j}(\lambda_j)$, $j = \alpha+1, \ldots, \beta$, contain the Jordan blocks of $\mathcal{A}$ corresponding to complex conjugate pairs of the non-real eigenvalues. Note that here the sizes of the Jordan blocks are $n_j/2$. If $n_j/2 \geq m+1$, then the diagonal of P contains a block $I_{n_j}^r$ with $n_j \geq 2m+2$. Now the same argument as above yields a contradiction, which shows that $n_j/2 \leq m$. ∎

As shown in the following example, the upper bound for the size of a real Jordan block of $\mathcal{A}$ in Theorem 2.3 can be attained.

EXAMPLE 2.4. The matrix

$$X := \frac{1}{\sqrt{2}} \left(\begin{bmatrix} I_m & & \\ & \sqrt{2} & \\ & & -I_m \end{bmatrix} + \begin{bmatrix} & & I_m^r \\ & 0 & \\ I_m^r & & \end{bmatrix} \right) \in \mathbb{R}^{2m+1,2m+1}$$

is nonsingular and satisfies $X^2 = I_{2m+1}$, i.e., $X = X^{-1}$. A tedious but straightforward computation shows that

$$X J_{2m+1}(0) X^{-1} = \frac{1}{2} \begin{bmatrix} J_m(0) + J_m(0)^H & \sqrt{2}e_m & J_m(0)I_m^r - I_m^r J_m(0) \\ \sqrt{2}e_m^H & 0 & -\sqrt{2}e_1^H \\ I_m^r J_m(0) - J_m(0) I_m^r & \sqrt{2}e_1 & J_m(0) + J_m(0)^H \end{bmatrix} =: \mathcal{A},$$

where $\mathcal{A} = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} \in \mathbb{R}^{2m+1,2m+1}$ has the blocks

$$A := \frac{1}{2} \begin{bmatrix} J_m(0) + J_m(0)^T & \sqrt{2}e_m \\ \sqrt{2}e_m^T & 0 \end{bmatrix},$$
$$B := \frac{1}{2} \begin{bmatrix} J_m(0)I_m^r - I_m^r J_m(0) & -\sqrt{2}e_1 \end{bmatrix},$$
$$C := \frac{1}{2}(J_m(0) + J_m(0)^T).$$

Here C is equal to the top $m \times m$ block of A, and thus the Cauchy interlacing theorem (see, e.g., [9, Theorem 4.3.17]) implies that the eigenvalues of A and C are strictly interlacing. Furthermore, the spectrum of C is given by $\sigma(C) = \left\{ \cos\left(\frac{\pi k}{m+1}\right) \mid k = 1, \ldots, m \right\}$; see, e.g., [9, p. 82].

Using Theorem 2.1 we can prove the following characterization of the largest possible non-diagonal part in the Jordan normal form of $\mathcal{A}$.

THEOREM 2.5. *If $\mathcal{A}$ is as in (1.2), then there exists a nonsingular matrix $X \in \mathbb{C}^{n+m,n+m}$ such that*

$$X^{-1} \mathcal{A} X = \begin{bmatrix} D & 0 \\ 0 & J \end{bmatrix},$$

where $D \in \mathbb{R}^{n+m-k,n+m-k}$ is diagonal and $J = \operatorname{diag}(J_{s_1}(\mu_1) \ldots J_{s_\ell}(\mu_\ell)) \in \mathbb{C}^{k,k}$ consists of Jordan blocks with $s_j \geq 2$, for $j = 1, \ldots, \ell$, and $\sum_{j=1}^{\ell} s_j = k \leq 3m$.

Proof. As in the proof of Theorem 2.3 we start with the decompositions $\mathcal{A} = X^{-1} J X$ and $\mathcal{J} = X^H P X$, where the matrices J and P satisfy (2.2) and (2.3), respectively. Without loss of

ETNA
Kent State University and
Johann Radon Institute (RICAM)

generality we can assume that $n_1 = \cdots = n_\gamma = 1$ for some $\gamma \leq \alpha$ (these are the 1×1 Jordan blocks corresponding to real eigenvalues of $\mathcal{A}$) and $n_{\gamma+1}, \ldots, n_\alpha \geq 2$ (these are the larger Jordan blocks corresponding to real eigenvalues). Moreover, we have $n_{\alpha+1}, \ldots, n_\beta \geq 2$.

We have to show that $\sum_{j=\gamma+1}^{\beta} n_j \leq 3m$. Since P and $\mathcal{J}$ are congruent and $\mathcal{J}$ has exactly m negative eigenvalues, we know that P must also have exactly m negative eigenvalues. Since the matrix $I_{n_j}^r$ has exactly $\lfloor \frac{n_j}{2} \rfloor$ eigenvalues -1 and $\lceil \frac{n_j}{2} \rceil$ eigenvalues $+1$, the number of negative eigenvalues of P is given by

$$m = \left(\sum_{j=1}^{\alpha} \delta_j \right) + \left(\sum_{j=\alpha+1}^{\beta} \frac{n_j}{2} \right), \quad \text{where} \quad \delta_j := \begin{cases} \lfloor \frac{n_j}{2} \rfloor & \text{if } \varepsilon_j = 1, \\ \lceil \frac{n_j}{2} \rceil & \text{if } \varepsilon_j = -1. \end{cases}$$

Note that $n_j \leq 2\lfloor \frac{n_j}{2} \rfloor + 1$, for $j = 1, \ldots, \beta$, and therefore

$$\sum_{j=\gamma+1}^{\beta} n_j \leq \sum_{j=\gamma+1}^{\beta} \left(2\lfloor \frac{n_j}{2} \rfloor + 1 \right) = 2 \left(\sum_{j=\gamma+1}^{\beta} \lfloor \frac{n_j}{2} \rfloor \right) + (\beta - \gamma) \leq 2m + (\beta - \gamma).$$

For $j = \gamma + 1, \ldots, \beta$ we have $n_j \geq 2$, and hence,

$$\beta - \gamma \leq \left(\sum_{j=\gamma+1}^{\alpha} \delta_j \right) + \left(\sum_{j=\alpha+1}^{\beta} \lfloor \frac{n_j}{2} \rfloor \right) \leq m,$$

which shows that $\sum_{j=\gamma+1}^{\beta} n_j \leq 3m$. ∎

A construction similar to the one in Example 2.4 can be used to show that the case $k = 3m$ in Theorem 2.5 can be attained; see also Example 3.8 below. Furthermore, a closer inspection of the proof of Theorem 2.5 shows that if $k = 3m$, then all eigenvalues of $\mathcal{A}$ have to be real and J consists solely of Jordan blocks of size three. (This is due to the fact that each complex conjugate pair of eigenvalues of $\mathcal{A}$ results in a diagonal block $I_{n_j}^r$ of even size in the matrix P.)

3. Conditions for a real spectrum and diagonalizability of $\mathcal{A}$. The goal of this section is to derive conditions for the blocks A, B, and C so that the matrix $\mathcal{A}$ in (1.2) has a real spectrum and is diagonalizable. Our derivations are based on the following two lemmas.

The first lemma is a special case of [14, Corollary 3.4]. We give a short proof for completeness.

LEMMA 3.1. *Let $M = M^T \in \mathbb{R}^{n,n}$ and $E \in \mathbb{R}^{n,n}$. Then every $\nu \in \sigma(M + E)$ satisfies*

$$\min_{\lambda \in \sigma(M)} |\lambda - \nu| \leq \|E\|_2 .$$

Proof. If ν is an eigenvalue of $M + E$ with corresponding eigenvector x, then

$$(M - \nu I_n)x = -Ex.$$

If $\nu \in \sigma(M)$, then trivially $\min_{\lambda \in \sigma(M)} |\nu - \lambda| = 0 \leq \|E\|_2$. If $\nu \notin \sigma(M)$, then $M - \nu I_n$ is nonsingular, and we obtain

$$x = -(M - \nu I_n)^{-1} Ex.$$

Taking 2-norms yields $\|x\|_2 \leq \left\|(M - \nu I_n)^{-1}\right\|_2 \|E\|_2 \|x\|_2$. Since M is symmetric, we obtain $\left\|(M - \nu I_n)^{-1}\right\|_2 = (\min_{\lambda \in \sigma(M)} |\lambda - \nu|)^{-1}$, which yields the desired inequality. ∎

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The second lemma was shown (under the assumption $A > 0$) in [2, Proposition 2.12]. It is also an immediate consequence of the more abstract result in [7, Theorem 4.2.4].

LEMMA 3.2. *If $\mathcal{A}$ as in (1.2) has a non-real eigenvalue, then any corresponding eigenvector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^{n+m}$ satisfies $\|x_1\|_2 = \|x_2\|_2$.*

We can now state and prove the first main result of this section.

THEOREM 3.3. *A matrix $\mathcal{A}$ as in (1.2) has a real spectrum if*

$$\|B\|_2 \le \frac{\gamma_{\mathcal{A}}}{2}, \quad \text{where} \quad \gamma_{\mathcal{A}} := \min_{\substack{\lambda \in \sigma(A) \\ \mu \in \sigma(C)}} |\lambda - \mu| .$$

Moreover, this bound is sharp in the sense that for every pair of matrices $A = A^T \in \mathbb{R}^{n,n}$ and $C = C^T \in \mathbb{R}^{m,m}$, there exists a $B \in \mathbb{R}^{m,n}$ with $\|B\|_2 > \gamma_A/2$ so that $\sigma(\mathcal{A})$ is not real. (In particular, if $\gamma_A = 0$, then there exists a $B \neq 0$ so that $\sigma(\mathcal{A})$ is not real.)

Proof. Suppose that $\mathcal{A}$ is as in (1.2) with $\|B\|_2 \le \gamma_{\mathcal{A}}/2$. We assume that $\mathcal{A}$ has a non-real eigenvalue ν, and we will derive a contradiction.

Splitting $\mathcal{A}$ in the form (1.3) and using Lemma 3.1 we see that

$$\min_{\eta \in \sigma(\operatorname{diag}(A,C))} |\eta - \nu| \le \left\| \begin{bmatrix} 0 & B^T \\ -B & 0 \end{bmatrix} \right\|_2 = \|B\|_2 ,$$

and therefore

$$\min_{\lambda \in \sigma(A)} |\lambda - \nu| \le \|B\|_2 \quad \text{or} \quad \min_{\mu \in \sigma(C)} |\mu - \nu| \le \|B\|_2 . \tag{3.1}$$

Let us assume that the first inequality holds, and denote

$$\widehat{\lambda} := \operatorname*{argmin}_{\lambda \in \sigma(A)} |\lambda - \nu| .$$

Then, by assumption,

$$\frac{\gamma_{\mathcal{A}}^2}{4} \ge \|B\|_2^2 \ge \left|\widehat{\lambda} - \nu\right|^2 = \left|\widehat{\lambda} - \operatorname{Re}(\nu)\right|^2 + \operatorname{Im}(\nu)^2,$$

and $\operatorname{Im}(\nu) \neq 0$ implies that $\left|\widehat{\lambda} - \operatorname{Re}(\nu)\right| < \gamma_{\mathcal{A}}/2$. Using the definition of $\gamma_{\mathcal{A}}$ we obtain

$$\gamma_{\mathcal{A}} \le \min_{\mu \in \sigma(C)} \left|\widehat{\lambda} - \mu\right| \le \min_{\mu \in \sigma(C)} \left(\left|\widehat{\lambda} - \operatorname{Re}(\nu)\right| + \left| \operatorname{Re}(\nu) - \mu \right| \right) < \frac{\gamma_{\mathcal{A}}}{2} + \min_{\mu \in \sigma(C)} |\mu - \nu| ,$$

which yields $\min_{\mu \in \sigma(C)} |\mu - \nu| > \gamma_{\mathcal{A}}/2$. Now let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be an eigenvector of $\mathcal{A}$ corresponding to the eigenvalue ν, i.e.,

$$\begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \nu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} . \tag{3.2}$$

The second block row can be rewritten as

$$-Bx_1 + (C - \nu I_m)x_2 = 0.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

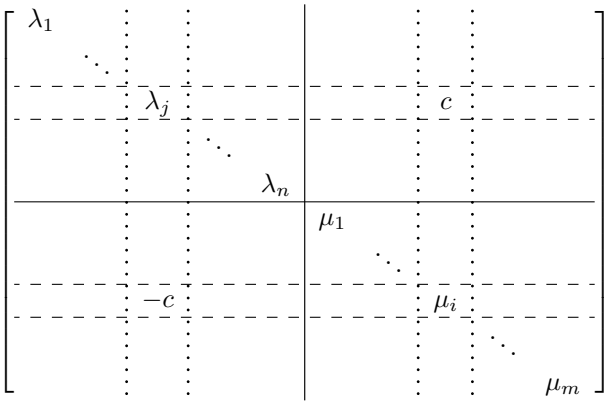
By Lemma 3.2 we know that $\|x_1\|_2 = \|x_2\|_2 > 0$, so without loss of generality we can assume that $\|x_1\|_2 = \|x_2\|_2 = 1$. Then we get

$$\|-Bx_1 + (C - \nu I_m)x_2\|_2 \geq \|(C - \nu I_m)x_2\|_2 - \|Bx_1\|_2 \geq \min_{\mu \in \sigma(C)} |\mu - \nu| - \frac{\gamma_{\mathcal{A}}}{2} > 0,$$

which is a contradiction.

A contradiction starting from the second inequality in (3.1) can be shown analogously by considering $\widehat{\mu} := \operatorname{argmin}_{\mu \in \sigma(C)} |\mu - \nu|$ and then the first block row of (3.2). Consequently, $\mathcal{A}$ cannot have a non-real eigenvalue.

In order to show that the bound for $\|B\|_2$ is sharp, let $A = A^T \in \mathbb{R}^{n,n}$ and $C = C^T \in \mathbb{R}^{m,m}$ be arbitrary. There exist orthogonal matrices $X_A \in \mathbb{R}^{n,n}$ and $X_C \in \mathbb{R}^{m,m}$ such that $X_A^T A X_A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $X_C^T C X_C = \operatorname{diag}(\mu_1, \ldots, \mu_m)$. Let $\lambda_j \in \sigma(A)$ and $\mu_i \in \sigma(C)$ be such that $|\lambda_j - \mu_i| = \gamma_{\mathcal{A}}$. Define $\widetilde{B} := [b_{kl}]$ with $b_{kl} = 0$ for $(k,l) \neq (i,j)$, and $b_{ij} = c$, where $c > \gamma_{\mathcal{A}}/2$. Then $B := X_A \widetilde{B} X_C$ satisfies $\|B\|_2 = \|\widetilde{B}\|_2 = c > \gamma_{\mathcal{A}}/2$. Furthermore, the eigenvalues of $\mathcal{A}$ coincide with the eigenvalues of the matrix



Since $\frac{(\lambda_j - \mu_i)^2}{4} = \frac{\gamma_{\mathcal{A}}^2}{4} < c^2$, the matrix $\mathcal{A}$ has the non-real eigenvalues $\frac{\lambda_j + \mu_i}{2} \pm \sqrt{\frac{\gamma_{\mathcal{A}}^2}{4} - c^2}$. □

It is clear from the splitting (1.3) that if in addition to the assumptions in Theorem 3.3 we have $A > 0$ and $C \geq 0$, then $\sigma(\mathcal{A}) \subset [0, \infty)$, i.e, $\mathcal{A}$ is (nonsymmetric) positive semidefinite. If $\mathcal{A}$ is also nonsingular, then $\sigma(\mathcal{A}) \subset (0, \infty)$, i.e., $\mathcal{A}$ is (nonsymmetric) positive definite.

Having discussed the realness of the eigenvalues of $\mathcal{A}$, we are now interested in conditions guaranteeing that $\mathcal{A}$ is diagonalizable.

LEMMA 3.4. *Let $\mathcal{A}$ be as in (1.2) with $\|B\|_2 < \gamma_{\mathcal{A}}/2$. Let $\nu \in \sigma(\mathcal{A})$ ($\subset \mathbb{R}$ by Theorem 3.3), and let $x \in \mathbb{R}^{n+m}$ be a corresponding eigenvector. Then exactly one of the following two cases holds:*

1. *If $\min_{\lambda \in \sigma(A)} |\lambda - \nu| < \gamma_{\mathcal{A}}/2$, then $x^T \mathcal{J} x > 0$.*
2. *If $\min_{\mu \in \sigma(C)} |\mu - \nu| < \gamma_{\mathcal{A}}/2$, then $x^T \mathcal{J} x < 0$.*

Proof. First note that since $\|B\|_2 < \gamma_{\mathcal{A}}/2$, the inequalities in (3.1) imply that

$$\text{(i)} \ \min_{\lambda \in \sigma(A)} |\lambda - \nu| < \frac{\gamma_{\mathcal{A}}}{2} \qquad \text{or} \qquad \text{(ii)} \ \min_{\mu \in \sigma(C)} |\mu - \nu| < \frac{\gamma_{\mathcal{A}}}{2}.$$

Only one of the cases can occur since we have

$$\gamma_A = \min_{\substack{\lambda \in \sigma(A) \\ \mu \in \sigma(C)}} |\lambda - \mu| \leq \min_{\substack{\lambda \in \sigma(A) \\ \mu \in \sigma(C)}} (|\lambda - \nu| + |\nu - \mu|) = \min_{\lambda \in \sigma(A)} |\lambda - \nu| + \min_{\mu \in \sigma(C)} |\mu - \nu|.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Assume that case (i) holds, and hence $\min_{\mu\in\sigma(C)}|\mu-\nu|>\gamma_{\mathcal{A}}/2$. Let $x=\begin{bmatrix}x_1\\x_2\end{bmatrix}$. Then the second block row of (3.2) yields

$$x_2=(C-\nu I_m)^{-1}Bx_1.$$

We obtain

$$\|x_2\|_2\le\frac{1}{\min_{\mu\in\sigma(C)}|\mu-\nu|}\|B\|_2\|x_1\|_2<\frac{2}{\gamma_{\mathcal{A}}}\frac{\gamma_{\mathcal{A}}}{2}\|x_1\|_2=\|x_1\|_2,$$

and hence

$$x^T\mathcal{J}x=\|x_1\|_2^2-\|x_2\|_2^2>0.$$

The proof for case (ii) follows analogously. □

We use Lemma 3.4 to prove the following result about the diagonalizability of $\mathcal{A}$.

THEOREM 3.5. *If $\mathcal{A}$ is as in (1.2) with $\|B\|_2<\gamma_{\mathcal{A}}/2$, then $\mathcal{A}$ is diagonalizable.*

Proof. First note that $\sigma(\mathcal{A})\subset\mathbb{R}$ by Theorem 3.3. We will show that $\mathcal{A}$ cannot have a Jordan block of size larger than one.

Let $\nu\in\sigma(\mathcal{A})$ and $y\in\ker((\nu I_{n+m}-\mathcal{A})^2)\subset\mathbb{R}^{n+m}$. Then it suffices to show that $y\in\ker(\nu I_{n+m}-\mathcal{A})$. The vector $x:=(\nu I_{n+m}-\mathcal{A})y\in\mathbb{R}^{n+m}$ satisfies

$$\begin{aligned}x^T\mathcal{J}x&=y^T(\nu I_{n+m}-\mathcal{A})^T\mathcal{J}(\nu I_{n+m}-\mathcal{A})y=y^T\mathcal{J}(\nu I_{n+m}-\mathcal{J}\mathcal{A}^T\mathcal{J})(\nu I_{n+m}-\mathcal{A})y\\&=y^T\mathcal{J}(\nu I_{n+m}-\mathcal{A})^2y=0,\end{aligned}$$

where we used that $\mathcal{A}$ is $\mathcal{J}$-selfadjoint. By Lemma 3.4, the vector x cannot be an eigenvector of $\mathcal{A}$ (since in that case $x^T\mathcal{J}x\neq 0$). But since $x\in\ker(\nu I_{n+m}-\mathcal{A})$, we must have $x=0$, and therefore $y\in\ker(\nu I_{n+m}-\mathcal{A})$. □

We point out that if $\mathcal{A}$ is as in (1.2) with $\|B\|_2<\gamma_{\mathcal{A}}/2$, then Lemma 3.4 can be used to show that $\mathcal{A}$ is even *stably diagonalizable* in the sense of [7, Section 9.2].

We will now illustrate Theorems 3.3 and 3.5 with several examples. The first example shows that if $\mathcal{A}$ is as in (1.2) with $\|B\|_2=\gamma_{\mathcal{A}}/2$, then $\mathcal{A}$ has a real spectrum by Theorem 3.3 but $\mathcal{A}$ may or may not be diagonalizable.

EXAMPLE 3.6. Consider the matrices

$$\mathcal{A}_1=\begin{bmatrix}2&0\\0&2\end{bmatrix}\quad\text{and}\quad\mathcal{A}_2=\begin{bmatrix}3&1\\-1&1\end{bmatrix}.$$

The matrix $\mathcal{A}_1$ is diagonalizable even though $\|B\|_2=\gamma_{\mathcal{A}_1}/2=0$. On the other hand, the matrix $\mathcal{A}_2$ with $\|B\|_2=\gamma_{\mathcal{A}_2}/2=1$ has the Jordan normal form

$$J=\begin{bmatrix}2&1\\0&2\end{bmatrix},$$

and hence $\mathcal{A}_2$ is not diagonalizable.

It is important to note that Theorems 3.3 and 3.5 for given symmetric matrices A and C give conditions so that $\mathcal{A}$ has a real spectrum and is diagonalizable, respectively, for *all possible matrices* B, as long as $\|B\|_2$ is small enough. As the following example shows, for some particular matrices B, the condition for $\|B\|_2$ is not strict.

EXAMPLE 3.7. Consider the matrix

$$\mathcal{A}=\begin{bmatrix}7&0&2\\0&0&0\\-2&0&2\end{bmatrix},\quad\text{where}\quad\|B\|_2=2>\frac{\gamma_{\mathcal{A}}}{2}=1.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Then $\sigma(\mathcal{A}) = \{0, 3, 6\}$, and hence $\mathcal{A}$ is diagonalizable with a real spectrum even though $\|B\|_2 > \gamma_{\mathcal{A}}/2$. By exchanging the last two rows and columns of $\mathcal{A}$ (which is a similarity transformation), we obtain the matrix

$$\widetilde{\mathcal{A}} = \begin{bmatrix} 7 & 2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} =: \begin{bmatrix} \mathcal{A}_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

For the matrix $\mathcal{A}_1$ we have $\|B\|_2 = 2 < \gamma_{\mathcal{A}_1}/2 = 5/2$, and hence Theorems 3.3 and 3.5 are applicable to $\mathcal{A}_1$, which also shows that $\mathcal{A}$ is diagonalizable with a real spectrum.

Suppose that $\mathcal{A}$ is as in (1.2) with $\|B\|_2 < \gamma_{\mathcal{A}}/2$, so that $\mathcal{A}$ has a real spectrum and is diagonalizable by Theorems 3.3 and 3.5. If we now increase the norm of the off-diagonal block by considering βB for $\beta > 1$, then eventually the resulting matrix $\mathcal{A}_\beta$ will have some non-real eigenvalues and/or will no longer be diagonalizable. As the following example illustrates, when considered as a function of (the increasing) β, the matrices $\mathcal{A}_\beta$ first form at least one Jordan block of size at least two, which then "splits" into complex conjugate pairs of eigenvalues. This observation also indicates that in general the formation of a Jordan block requires a "normwise smaller" off-diagonal block than the formation of a complex conjugate pair of eigenvalues.

EXAMPLE 3.8. For the matrix

$$\mathcal{A}_\beta = \begin{bmatrix} 2 & 0 & -\beta \\ 0 & -2 & \beta \\ \beta & -\beta & 0 \end{bmatrix}, \quad \text{where} \quad \|B\|_2 = \sqrt{2}\beta \quad \text{and} \quad \frac{\gamma_{\mathcal{A}_\beta}}{2} = 1,$$

we have $\sigma(\mathcal{A}_\beta) = \{0, \sqrt{4 - 2\beta^2}, -\sqrt{4 - 2\beta^2}\}$, and hence $\sigma(\mathcal{A}_\beta) \subset \mathbb{R}$ for all $\beta \leq \sqrt{2}$, and $\mathcal{A}_\beta$ is diagonalizable for all $\beta < \sqrt{2}$. Note that, in contrast to Example 3.7, there exists no simultaneous permutation of rows and columns such that $\mathcal{A}_\beta$ is a block diagonal matrix.

For any $\beta > \sqrt{2}$, the matrix $\mathcal{A}_\beta$ has non-real eigenvalues, and for $\beta = \sqrt{2}$, the matrix

$$\mathcal{A}_{\sqrt{2}} = \begin{bmatrix} 2 & 0 & -\sqrt{2} \\ 0 & -2 & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{bmatrix}$$

has the Jordan normal form

$$J_{\sqrt{2}} = \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix},$$

which also gives an example of a largest possible Jordan block corresponding to a real eigenvalue; see Theorem 2.3.

In the next two examples we illustrate our results using some larger-scale matrices from a finite element discretization of a Stokes model problem and the time discretization of a mass-spring-damper system. The main goal in the first example is to show the intriguing behavior that has to be expected when analyzing the spectrum of nonsymmetric saddle point matrices. The second example shows that in some problems the conditions for diagonalizability with real eigenvalues are satisfied naturally for certain parameter choices (here the length of a time step).

EXAMPLE 3.9. We consider the classical example of a steady horizontal flow in a channel, which is driven by a pressure difference. Following [5, Section 5], this problem is modeled by

ETNA
Kent State University and
Johann Radon Institute (RICAM)

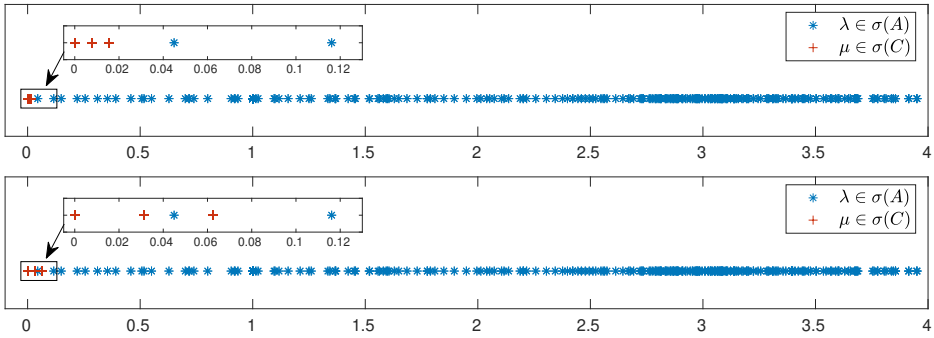


FIG. 3.1. *The spectra of A and C from Example 3.9 for $\beta = \frac{1}{4}$ (top) and $\beta = 1$ (bottom).*

the equations

$$\begin{aligned} -\nabla^2 u + \nabla p &= 0 && \text{in } \Omega, \\ \nabla \cdot u &= 0 && \text{in } \Omega, \\ u &= [1 - y^2,\ 0]^T && \text{on } \partial\Omega, \end{aligned} \tag{3.3}$$

where $\Omega = (-1, 1)^2$. The discretization of (3.3) with stabilized finite elements is one of the test problems implemented in IFISS [13]. Running the *channel domain* example in the IFISS `stokes_testproblem` with the default parameters[1] leads to a linear algebraic system with the matrix

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ -B & \beta C \end{bmatrix} \in \mathbb{R}^{834,834},$$

where $\beta > 0$ is a stabilization parameter. The default parameter is $\beta = 1/4$, and we also use $\beta = 1$. For these two parameters we obtain the following values, computed by MATLAB's `eig` and `norm` functions:

TABLE 3.1
Numerically computed values for the matrices in our model problem.

β	$\lambda_{\min}(\beta C)$	$\lambda_{\max}(\beta C)$	$\lambda_{\min}(A)$	$\lambda_{\max}(A)$	$\|B\|_2$	$\gamma_{\mathcal{A}}/2$
$\frac{1}{4}$	0	0.0156	0.0449	3.9515	0.2477	0.0147
1	0	0.0625	0.0449	3.9515	0.2477	0.0068

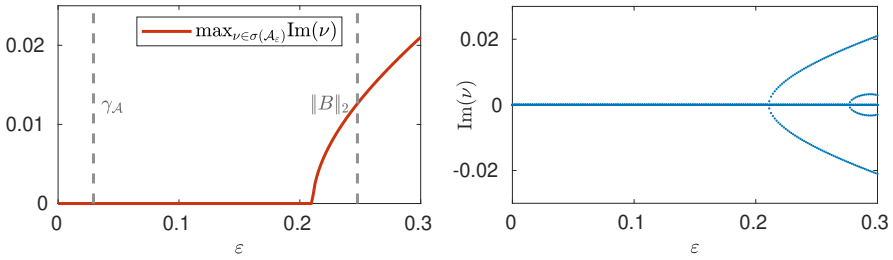
We observe that $\sigma(A)$ and $\sigma(\beta C)$ are strictly separated for $\beta = 1/4$ and that they interlace for $\beta = 1$; see also Figure 3.1. In both cases $\|B\|_2 > \gamma_{\mathcal{A}}/2$, and hence the results of Section 3 are not applicable. Indeed, a computation of $\sigma(\mathcal{A})$ with MATLAB's `eig` function shows that $\mathcal{A}$ has non-real eigenvalues for both values of β.

In order to make the results from Section 3 applicable, we scale the off-diagonal blocks, i.e., we consider the family of matrices

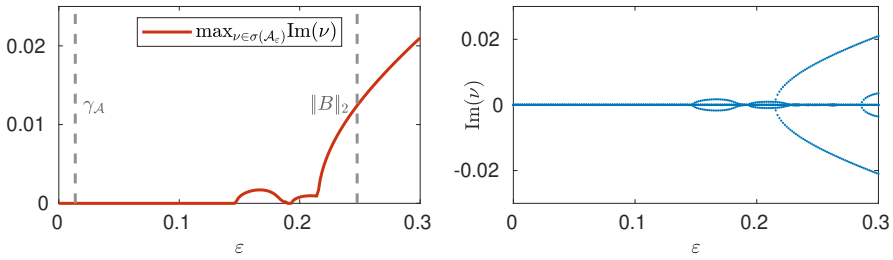
$$\mathcal{A}_\varepsilon = \begin{bmatrix} A & \frac{\varepsilon}{\|B\|_2} B^T \\ -\frac{\varepsilon}{\|B\|_2} B & \beta C \end{bmatrix}, \quad \varepsilon \geq 0.$$

[1]The default parameters for the *channel domain* test problem are: channel domain: unity; grid parameter: 4 (16x16 grid); Q1-P0-elements; stabilization parameter: $\frac{1}{4}$; uniform streamlines. All numerical results were computed using MATLAB version R2023a on an Intel i5-1235U and 16 GB RAM.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $\beta = \frac{1}{4}$. (b) $\beta = \frac{1}{4}$.



(c) $\beta = 1$. (d) $\beta = 1$.

FIG. 3.2. *Left: maximal imaginary part of the eigenvalues of $\mathcal{A}_\varepsilon$ for $\beta = \frac{1}{4}$ (top) and $\beta = 1$ (bottom). Right: imaginary part of the eigenvalues of $\mathcal{A}_\varepsilon$ for $\beta = \frac{1}{4}$ (top) and $\beta = 1$ (bottom).*

Note that the off-diagonal blocks of $\mathcal{A}_\varepsilon$ have the 2-norm ε, so that $\mathcal{A}_\varepsilon$ is diagonalizable with a real spectrum when $\varepsilon < \gamma_{\mathcal{A}}/2$.

In Figure 3.2 we display the maximum imaginary part of $\sigma(\mathcal{A}_\varepsilon)$ for increasing values of ε. We see that for both values of β, the eigenvalues of $\mathcal{A}_\varepsilon$ stay real even when $\varepsilon \gg \gamma_{\mathcal{A}}/2$. Another interesting behavior can be observed for the parameter $\beta = 1$. The spectrum of $\mathcal{A}_\varepsilon$ becomes non-real at approximately $\varepsilon = 0.15$. For a small interval of slightly larger values (approximately starting at $\varepsilon = 0.19$), all eigenvalues become real again, followed by a non-real spectrum for all larger values of ε.

EXAMPLE 3.10. We consider a mass-spring-damper (MSD) system, which is a standard example in model order reduction and also has been used, e.g., in [8, 11]. A detailed derivation and physical interpretation of MSD systems can be found, e.g., in [16].

We consider $g \geq 2$ masses which are connected in the following way (see Figure 3.3): For $i = 1, \ldots, g-1$, the mass m_i is attached to the mass m_{i+1} by a spring with spring coefficient $k_{i,i+1}$ and a damper with damping coefficient $d_{i,i+1}$. Additionally, for $i = 1, \ldots, g$, the mass m_i is connected to the ground by a spring and a damper with spring and damping coefficients κ_i and δ_i, respectively. The vibration of this system is described by a second-order differential algebraic equation (DAE) of the form

$$M\ddot{x} + D\dot{x} + Kx = f,$$

where the entries of x correspond to the displacement of the masses and f is an external force. The matrices M,K, and D corresponding to the masses, the stiffness, and the damping are given by

ETNA
Kent State University and
Johann Radon Institute (RICAM)

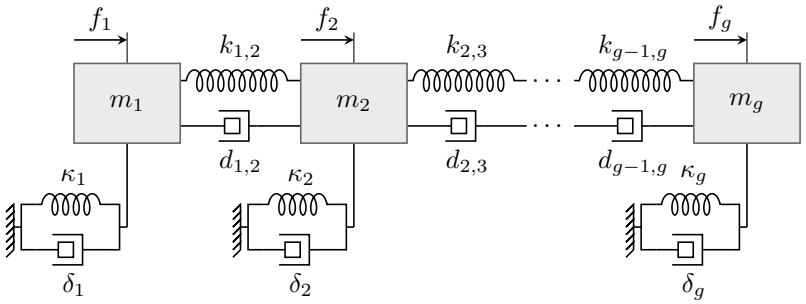


FIG. 3.3. *Illustration of a mass-spring-damper system.*

$$
\begin{aligned}
M &= \operatorname{diag}(m_1, \ldots, m_g) \in \mathbb{R}^{g,g},\\
D &= \operatorname{tridiag}(-d_{i-1,i},\ d_{i-1,i} + d_{i,i+1} + \delta_i,\ -d_{i,i+1}) \in \mathbb{R}^{g,g},\\
K &= \operatorname{tridiag}(-k_{i-1,i},\ k_{i-1,i} + k_{i,i+1} + \kappa_i,\ -k_{i,i+1}) \in \mathbb{R}^{g,g},
\end{aligned}
$$

where we set $d_{0,1} = d_{g,g+1} = k_{0,1} = k_{g,g+1} = 0$. We assume that all parameters $m_i, d_{i-1,i}, k_{i-1,i}, \kappa_i, \delta_i > 0$ and hence that the matrices are symmetric positive definite. The first-order formulation of the DAE is given by

$$
\begin{bmatrix} M & 0 \\ 0 & K \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} -D & -K \\ K & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} + \begin{bmatrix} f \\ 0 \end{bmatrix},
$$

where we multiplied the second block row by the matrix K. A time discretization using the implicit midpoint rule with uniform time steps t_k, $k = 0, 1, 2, \ldots$, of length $\tau > 0$ yields a sequence of linear algebraic systems of the form

$$
\mathcal{A} \begin{bmatrix} v_{k+1} \\ p_{k+1} \end{bmatrix} = b(v_k, p_k, \tau, t_k), \quad \text{where} \quad \mathcal{A} = \begin{bmatrix} M + \frac{\tau}{2} D & \frac{\tau}{2} K \\ -\frac{\tau}{2} K & K \end{bmatrix}.
$$

Thus, each time step requires to solve a linear algebraic system with a nonsymmetric positive definite saddle point matrix $\mathcal{A}$ of the form (1.2), where $M + \frac{\tau}{2} D > 0$ and $K > 0$. The 2-norm of the off-diagonal blocks of $\mathcal{A}$ is given by $\frac{\tau}{2} \|K\|_2$. This means that for any sufficiently small step length $\tau > 0$, it is guaranteed that the nonsymmetric matrix $\mathcal{A}$ will be diagonalizable with real and positive eigenvalues (as long as the spectra of $M + \frac{\tau}{2} D$ and K do not intersect).

For a numerical illustration we consider $g = 100$, $\tau = 10^{-3}$, and the remaining parameters as in Table 3.2. For these parameters, the spectra of the matrices $M + \frac{\tau}{2} D$ and K interlace (but do not intersect), which can be seen in Figure 3.4. Using MATLAB's `eig` and `norm` functions we compute $\gamma_{\mathcal{A}}/2 \approx 0.0118$ and $\frac{\tau}{2} \|K\|_2 \approx 0.005$, so that $\gamma_{\mathcal{A}}/2 < \frac{\tau}{2} \|K\|_2$. Theorems 3.3 and 3.5 now imply that the matrix $\mathcal{A}$ is diagonalizable with real and positive eigenvalues. (These are contained in the interval $[2, 10]$.)

4. Finding an inner-product matrix for $\mathcal{A}$. As shown in [10, Lemma 2.1], for any (real) polynomial p, the matrix $\mathcal{J}p(\mathcal{A})$ is symmetric, and the matrix $\mathcal{A}$ is $\mathcal{J}p(\mathcal{A})$-selfadjoint. Thus, the condition $\mathcal{J}p(\mathcal{A}) > 0$ is sufficient for $\mathcal{A}$ to be diagonalizable with a real spectrum; cf. [10, Corollary 3.2]. Necessary and sufficient conditions for positive definiteness of the particular matrix

$$
\mathcal{M}(\gamma) = \begin{bmatrix} A - \gamma I_n & B^T \\ B & \gamma I_m - C \end{bmatrix} = \mathcal{J}p(\mathcal{A}), \qquad \text{where} \quad p(z) := z - \gamma \quad \text{(see (1.4))},
$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 3.2
Parameters for the MSD system.

$m_i =$	$k_{i,i+1} =$	$\kappa_i =$	$d_{i,i+1} =$	$\delta_i =$
$\begin{cases} 4 & i \le 20 \\ 5 & 20 < i \le 40 \\ 6 & 40 < i \le 60 \\ 7 & 60 < i \le 80 \\ 8 & 80 < i \end{cases}$	2	$\begin{cases} 4 & i = 1, g \\ 2 & \text{else} \end{cases}$	5	$\begin{cases} 10 & i = 1, g \\ 5 & \text{else} \end{cases}$

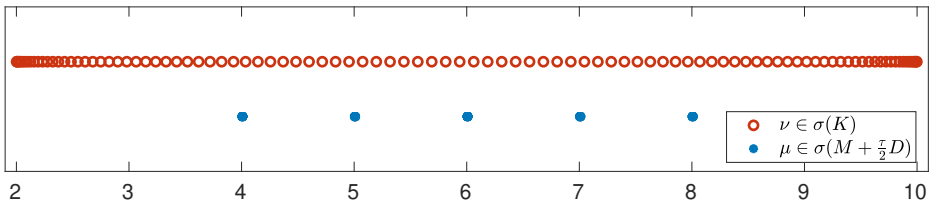


FIG. 3.4. *The spectra of $M + \frac{\tau}{2}D$ and K from Example 3.10.*

were derived in [10, Theorem 2.2]; see also the earlier and closely related investigation in [2, Section 3], which assumes that $C = 0$. One of the necessary conditions for $\mathcal{M}(\gamma) > 0$ is $\lambda_{\min}(A) > \gamma > \lambda_{\max}(C)$, and hence any further condition on $\mathcal{A}$ derived in this way requires that $\lambda_{\min}(A) > \lambda_{\max}(C)$. Clearly, the conditions in Theorems 3.3 and 3.5 are more general since they show that $\mathcal{A}$ can be diagonalizable with a real spectrum even when $\sigma(A)$ and $\sigma(C)$ interlace.

The matrix $\mathcal{M}(\gamma)$ in [10] and the corresponding matrix in [2] were given without a strategy for generalization. We will now explain how to obtain a polynomial p with $\mathcal{J}p(\mathcal{A}) > 0$ and hence an inner product in which $\mathcal{A}$ is selfadjoint, for the case of interlacing $\sigma(A)$ and $\sigma(C)$. Suppose that $\mathcal{A}$ is as in (1.2) with $\|B\|_2 < \gamma_A/2$ so that $\mathcal{A}$ is diagonalizable with a real spectrum. Let $\nu_1, \ldots, \nu_k \in \mathbb{R}$ be the $k \le n + m$ distinct eigenvalues of $\mathcal{A}$. Then every $y \in \mathbb{R}^{n+m}$ can be written as

$$y = \sum_{j=1}^{k} \alpha_j x_j,$$

where $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ and $x_j \in \ker(\nu_j I_{n+m} - \mathcal{A})$, for $j = 1, \ldots, k$. By [7, Theorem 4.2.4], the vectors $x_1, \ldots, x_k$ are pairwise orthogonal in the indefinite $\mathcal{J}$-inner product. Therefore, for any polynomial p,

$$y^T \mathcal{J} p(\mathcal{A}) y = \sum_{j=1}^{k} \alpha_j^2 p(\nu_j) x_j^T \mathcal{J} x_j.$$

By Lemma 3.4 we have either $x_j^T \mathcal{J} x_j > 0$ or $x_j^T \mathcal{J} x_j < 0$, depending on whether

$$\text{(i)} \ \min_{\lambda \in \sigma(A)} |\lambda - \nu_j| < \frac{\gamma_{\mathcal{A}}}{2} \qquad \text{or} \qquad \text{(ii)} \ \min_{\mu \in \sigma(C)} |\mu - \nu_j| < \frac{\gamma_{\mathcal{A}}}{2},$$

respectively. Thus, if we choose a polynomial p such that

$$p(\nu_j) > 0 \quad \text{if } \nu_j \text{ satisfies (i)} \qquad \text{and} \qquad p(\nu_j) < 0 \quad \text{if } \nu_j \text{ satisfies (ii)}, \tag{4.1}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

then $y^T \mathcal{J} p(\mathcal{A}) y > 0$ holds for any nonzero $y \in \mathbb{R}^{n+m}$, i.e., we have $\mathcal{J} p(\mathcal{A}) > 0$.

The essential observation to be made from (4.1) for the construction of a polynomial p that satisfies $\mathcal{J} p(\mathcal{A}) > 0$ is that this polynomial must separate $\sigma(A)$ and $\sigma(C)$ in the sense that it is positive close to $\sigma(A)$ and negative close to $\sigma(C)$. If the two spectra are strictly separated with $\lambda_{\min}(A) > \lambda_{\max}(C)$, as assumed in [10], then a polynomial of degree one that has its only root between $\lambda_{\max}(C)$ and $\lambda_{\min}(A)$ works. This is just the result from [10, Corollary 3.1], where it is suggested to use

$$p(z) = z - \widehat{\gamma} \quad \text{with} \quad \widehat{\gamma} = \frac{\lambda_{\min}(A) + \lambda_{\max}(C)}{2}.$$

In this notation, an inner-product matrix $\mathcal{J} p(\mathcal{A})$ with $p(z) = z - \lambda_{\min}(A)/2$ was considered in [2, Proposition 3.1].

Here we can also deal with interlacing $\sigma(A)$ and $\sigma(C)$. For example, suppose that the eigenvalues λ_j and μ_j of A and C, respectively, are contained in three interlacing and non-intersecting intervals

$$[\lambda_n, \lambda_k], \quad [\mu_m, \mu_1], \quad [\lambda_{k+1}, \lambda_1],$$

where $\lambda_n \le \lambda_k < \mu_m \le \mu_1 < \lambda_{k+1} \le \lambda_1$, for some $k \in \{1, \dots, n-1\}$. Then the polynomial

$$p_2(z) = \left(z - \left(\mu_m - \frac{\gamma_A}{2}\right)\right)\left(z - \left(\mu_1 + \frac{\gamma_A}{2}\right)\right)$$

is negative (exactly) on the open interval $(\mu_m - \gamma_A/2, \mu_1 + \gamma_A/2)$. Hence it satisfies the properties stated in (4.1), which implies that $\mathcal{J} p_2(\mathcal{A}) > 0$.

In general, if the eigenvalues of C are contained in $q \ge 1$ intervals that interlace the eigenvalues of A, i.e.,

$$[\lambda_n, \lambda_{r_q+1}], \quad [\mu_m, \mu_{\ell_{q-1}+1}], \quad \dots, \quad [\mu_{\ell_2}, \mu_{\ell_1+1}], \quad [\lambda_{r_2}, \lambda_{r_1+1}], \quad [\mu_{\ell_1}, \mu_1], \quad [\lambda_{r_1}, \lambda_1],$$

then the polynomial

$$p_{2q}(z) = \prod_{j=1}^{q} \left(z - \left(\mu_{\ell_j} - \frac{\gamma_{\mathcal{A}}}{2}\right)\right)\left(z - \left(\mu_{\ell_{j-1}+1} + \frac{\gamma_{\mathcal{A}}}{2}\right)\right), \quad \text{where } \ell_q = m \text{ and } \ell_0 = 1,$$

satisfies the properties in (4.1) and hence leads to $\mathcal{J} p_{2q}(\mathcal{A}) > 0$.

The following figure displays a sketch of this polynomial for $q = 2$:

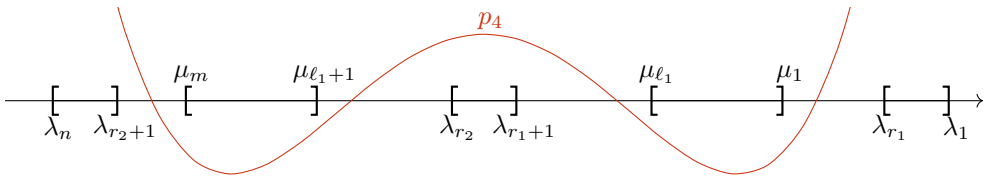


FIG. 4.1. *Sketch of the polynomial p_4.*

Depending on the number and ordering of these intervals, there may be polynomials of smaller degree than $2q$ that also separate $\sigma(A)$ and $\sigma(C)$ in the sense of (4.1). For example, if $\lambda_{r_2} = \lambda_n$ in the previous sketch, then the spectra can be separated by a polynomial of degree 3 (instead of 4); see Figure 4.2.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

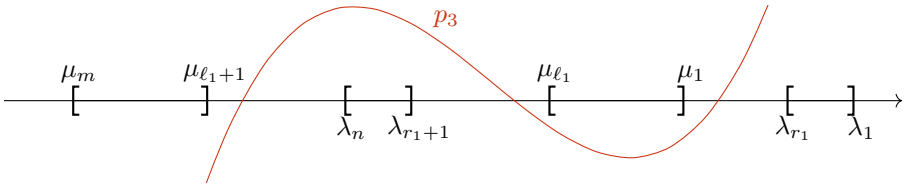


FIG. 4.2. *Sketch of a polynomial of degree* 3 *which satisfies* (4.1).

Note that the situation illustrated in Figure 4.2 occurs for the matrix $\mathcal{A}$ in Example 3.9 with $\beta = 1$; see the bottom plot in Figure 3.1. We summarize the previous considerations in the following corollary.

COROLLARY 4.1. *Let $\mathcal{A}$ be as in* (1.2) *with $\|B\|_2 < \gamma_{\mathcal{A}}/2$, and suppose that the eigenvalues of C are contained in $q \geq 1$ intervals that interlace the eigenvalues of A. Then there exists a polynomial p of degree at most $2q$ satisfying* (4.1) *and hence $\mathcal{J}p(\mathcal{A}) > 0$.*

We point out that even when only considering separating polynomials of smallest possible degree, these polynomials are not uniquely determined by (4.1) since the zeros can be placed anywhere between the spectral intervals as long as their distances to the respective eigenvalues of A and C is at least $\gamma_{\mathcal{A}}/2$.

5. Concluding remarks. We have analyzed spectral properties of nonsymmetric saddle point matrices $\mathcal{A}$ of the form (1.2). In addition to giving a complete analysis of the largest possible Jordan block sizes of $\mathcal{A}$ that can occur, we have generalized the previously known conditions for a real spectrum and diagonalizability of $\mathcal{A}$. Moreover, we have extended the approach in [2, 10] to find a nonstandard inner product in which $\mathcal{A}$ is selfadjoint.

For any matrix $\mathcal{A}$ that is diagonalizable with a real spectrum, our construction yields a symmetric matrix $\mathcal{J}p(\mathcal{A}) > 0$ with respect to which $\mathcal{A}$ is selfadjoint and hence a conjugate gradient method for solving linear algebraic systems with $\mathcal{A}$. It is clear that this is a more theoretical than practical result since the explicit construction of a polynomial that satisfies the conditions in (4.1) requires rather precise knowledge about $\sigma(A)$ and $\sigma(C)$, and the degree of such a polynomial is usually twice the number of the eigenvalue intervals of $\sigma(C)$ which interlace $\sigma(A)$.

REFERENCES

[1] M. BENZI, G. H. GOLUB, AND J. LIESEN, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137.
[2] M. BENZI AND V. SIMONCINI, *On the eigenvalues of a class of saddle point matrices*, Numer. Math., 103 (2006), pp. 173–196.
[3] L. BERGAMASCHI, *On eigenvalue distribution of constraint-preconditioned symmetric saddle point matrices*, Numer. Linear Algebra Appl., 19 (2012), pp. 754–772.
[4] J. H. BRAMBLE AND J. E. PASCIAK, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50 (1988), pp. 1–17.
[5] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, 2nd ed., Oxford University Press, Oxford, 2014.
[6] B. FISCHER, A. RAMAGE, D. J. SILVESTER, AND A. J. WATHEN, *Minimum residual methods for augmented systems*, BIT Numer. Math., 38 (1998), pp. 527–543.
[7] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Indefinite Linear Algebra and Applications*, Birkhäuser, Basel, 2005.
[8] C. GÜDÜCÜ, J. LIESEN, V. MEHRMANN, AND D. B. SZYLD, *On non-Hermitian positive (semi)definite linear algebraic systems arising from dissipative Hamiltonian DAEs*, SIAM J. Sci. Comput., 44 (2022), pp. A2871–A2894.
[9] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, 2nd ed. , Cambridge University Press, Cambridge, 2013.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[10] J. LIESEN AND B. N. PARLETT, *On nonsymmetric saddle point matrices that allow conjugate gradient iterations*, Numer. Math., 108 (2008), pp. 605–624.
[11] V. MEHRMANN AND T. STYKEL, *Balanced truncation model reduction for large-scale systems in descriptor form*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2005, pp. 83–115.
[12] J. SCHÖBERL AND W. ZULEHNER, *Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 752–773.
[13] D. SILVESTER, H. ELMAN, AND A. RAMAGE, *Incompressible Flow and Iterative Solver Software (IFISS) version 3.6*, Software Package, August 2021.
[14] G. W. STEWART AND J. G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[15] M. STOLL AND A. WATHEN, *Combination preconditioning and the Bramble-Pasciak$^+$ preconditioner*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 582–608.
[16] K. VESELIĆ, *Damped Oscillations of Linear Systems*, Springer, Heidelberg, 2011.