ON THE DENSITY OF SEMISIMPLE MATRICES IN INDEFINITE SCALAR PRODUCT SPACES*

RALPH J. DE LA CRUZ[†] AND PHILIP SALTENBERGER[‡]

Abstract. For an indefinite scalar product $[x,y]_B = x^H B y$ for $B = \pm B^H \in \mathbf{Gl}_n(\mathbb{C})$ on $\mathbb{C}^n \times \mathbb{C}^n$, it is shown that the set of diagonalizable matrices is dense in the set of all B-normal matrices. The analogous statement is also proven for the sets of B-selfadjoint, B-skewadjoint and B-unitary matrices.

Key words. Indefinite scalar products, Sesquilinear forms, Diagonalization, Density, Normal matrices, Selfadjoint matrices, Skewadjoint matrices, Unitary matrices.

AMS subject classifications. 15A63, 15A21, 11E39, 11R45.

1. Introduction. Whenever $B \in \mathbf{Gl}_n(\mathbb{C})$ is some arbitrary (nonsingular) matrix, the function

$$[\cdot, \cdot]_B : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \ (x, y) \mapsto x^H B y \quad (x^H := \overline{x}^T),$$

defines a nondegenerate sesquilinear form on $\mathbb{C}^n \times \mathbb{C}^n$. In case B is not necessarily Hermitian positive definite, such forms are also referred to as indefinite scalar products [13, Section 2]. In this work, we particularly consider such products for the cases $B = B^H$, $B = -B^H$ (with B^H denoting the conjugate transpose of B) and $B^H B = I_n$ (i.e. B is unitary). Indefinite scalar products with $B = \pm B^H$ are called orthosymmetric while those with $B^H = B^{-1}$ are called unitary [13, Definitions 3.1 and 4.1].

Several classes of matrices $A \in \mathbf{M}_n(\mathbb{C})$ are naturally related to $[\cdot, \cdot]_B$:

1. A matrix $J \in \mathbf{M}_n(\mathbb{C})$ is called B-selfadjoint if $[Jx, y]_B = [x, Jy]_B$ holds for all $x, y \in \mathbb{C}^n$. It follows that $x^H J^H B y = x^H B J y$ holds for all $x, y \in \mathbb{C}^n$ if J is B-selfadjoint. This means we have $J^H B = BJ$, that is, $J = B^{-1} J^H B$. The set of B-selfadjoint matrices is denoted by $\mathbb{J}(B)$.

2. A matrix $L \in \mathbf{M}_n(\mathbb{C})$ is called B-skewadjoint if $[Lx, y]_B = [x, -Ly]_B$ holds for all $x, y \in \mathbb{C}^n$. It follows from this equation as above that L is B-skewadjoint if and only if $-L = B^{-1} L^H B$. The set of all B-skewadjoint matrices is denoted by $\mathbb{L}(B)$.

3. A matrix $G \in \mathbf{M}_n(\mathbb{C})$ is called B-unitary if $[Gx, Gy]_B = [x, y]_B$ holds for all $x, y \in \mathbb{C}^n$. This means that $x^H G^H B G y = x^H B y$ has to hold for all $x, y \in \mathbb{C}^n$ and implies that $G^H B G = B$. The set of all B-unitary matrices is denoted by $\mathbb{G}(B)$.

4. A matrix $N \in \mathbf{M}_n(\mathbb{C})$ is called B-normal if $N B^{-1} N^H B = B^{-1} N^H B N$ holds. The set of all B-normal matrices is denoted by $\mathcal{N}(B)$.

Notice that the set $\mathcal{N}(B)$ of B-normal matrices includes the sets $\mathbb{J}(B), \mathbb{L}(B)$ and $\mathbb{G}(B)$. A great many of problems in control systems theory, matrix equations or differential equations involve matrices from

the classes defined above (see, e.g., [11, 13] and the references therein). The monograph [4] contains a comprehensive treatment and applications for the Hermitian case $B = B^H$.

Assume that $B = I_n$ is the $n \times n$ identity matrix. Then the sets of B-selfadjoint, B-skewadjoint, B-unitary and B-normal matrices $A \in \mathbf{M}_n(\mathbb{C})$ coincide with the sets of Hermitian $(A = A^H)$, skew-Hermitian $(A = -A^H)$, unitary $(A^H A = I_n)$ and normal $(AA^H = A^H A)$ matrices. It is a well-known fact that every matrix belonging to any of these four sets of matrices is semisimple [7]. Furthermore, any normal matrix $A \in \mathcal{N}(I_n)$ is unitarily diagonalizable. In case $B \neq I_n$, the situation is different: for instance, consider the 4×4 matrix $\Sigma_{2,2} \in \mathbf{M}_4(\mathbb{R})$ (which is Hermitian but not positive definite) and the matrix $N \in \mathbf{M}_4(\mathbb{C})$:

$$(1.1) \qquad N = \begin{bmatrix} 0.5 & 0 & 0 & -0.5 \\ 0 & -0.5+i & -0.5 & 0 \\ 0 & 0.5 & 0.5+i & 0 \\ 0.5 & 0 & 0 & 1.5 \end{bmatrix}, \qquad \Sigma_{2,2} = \begin{bmatrix} -1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}.$$

The matrix N is $\Sigma_{2,2}$-normal (but neither $\Sigma_{2,2}$-skew/selfadjoint nor $\Sigma_{2,2}$-unitary) and not semisimple. Its Jordan form consists of two Jordan blocks (both size 2×2) for the eigenvalues 1 and i. In general, a matrix $A \in \mathbf{M}_n(\mathbb{C})$ belonging to $\mathcal{N}(B)$ for some arbitrary (skew)-Hermitian matrix $B \in \mathbf{Gl}_n(\mathbb{C})$ need not be semisimple (the same is true for the subsets $\mathbb{J}(B), \mathbb{L}(B)$ and $\mathbb{G}(B)$). Thus, given some (skew)-Hermitian $B \in \mathbf{Gl}_n(\mathbb{C})$, the following question arises: 'How large is the subset of semisimple matrices in $\mathcal{N}(B)$?' The answer to this question is simply 'every matrix is semisimple' in case $B = I_n$ (and whenever B is Hermitian positive definite), but seems to be unknown for general indefinite or skew-Hermitian B.

In this work, we consider the question raised above from a topological point of view. Recall that $\mathbf{M}_n(\mathbb{C})$ can be considered as a topological space with basis $B_R(A) = \{A' \in \mathbf{M}_n(\mathbb{C}) : \|A - A'\| < R\}$ for $A \in \mathbf{M}_n(\mathbb{C})$ and $R \in \mathbb{R}, R > 0$ (see, e.g. [8, Section 11.2]). The sets $\mathbb{J}(B), \mathbb{L}(B), \mathbb{G}(B)$ and $\mathcal{N}(B)$ can thus all be interpreted as topological spaces on their own equipped with the subspace topology [1, Section 1.5]. For instance, a subset $\mathcal{S} \subset \mathcal{N}(B)$ is open (in this subspace topology) if $\mathcal{S}$ is the intersection of $\mathcal{N}(B)$ with some open subset of $\mathbf{M}_n(\mathbb{C})$. If a property does not hold for all elements in a topological space, it is usually reasonable to ask whether it holds for a dense subset. Thus, a second question, in this context naturally related to the first one, is 'As not all matrices in $\mathcal{N}(B)$ are semisimple, is the subset of semisimple matrices in $\mathcal{N}(B)$ at least dense?' Since the sets $\mathbb{J}(B), \mathbb{L}(B), \mathbb{G}(B)$ and $\mathcal{N}(B)$ are defined by matrix equations, they are (topologically) closed subsets (for instance, $N \in \mathbf{M}_n(\mathbb{C})$ is B-normal if and only if it satisfies the equation $NB^{-1}N^H B = B^{-1}N^H BN$). Consequently, although the set of semisimple matrices is dense in $\mathbf{M}_n(\mathbb{C})$ [12, Cor. 7.3.3], there is no direct reasoning why this fact should carry over to $\mathcal{N}(B)$. In fact, a closed subset of $\mathbf{M}_n(\mathbb{C})$ need not contain any diagonalizable matrix at all (see [8] for an example). This work is devoted to the second question raised above and gives a positive answer. Our main result can be stated as follows:

> **Main result.** Let $B = \pm B^H \in \mathbf{Gl}_n(\mathbb{C})$. Then the set of diagonalizable matrices is dense in $\mathcal{N}(B)$.

The proof of the density result for $\mathcal{N}(B)$ relies on observations on symmetric polynomials and 1-regularity of matrices (see Section 2.1). This approach is also applicable to provide direct proofs for the density of semisimple matrices in the sets $\mathbb{J}(B), \mathbb{L}(B)$ and $\mathbb{G}(B)$. For these three classes of matrices, the density result of semisimple matrices is not surprising, since Jordan-like canonical forms are available for these matrices, see [4, Section 5], [6]. The normal forms allow the explicit construction of arbitrary small structure-preserving perturbations (by adding ε to the corners of the Jordan blocks) such that the new matrix becomes semisimple (in particular, see Remark 3.4). This approach immediately yields the density result for matrices in $\mathbb{J}(B)$,

$\mathbb{L}(B)$ and $\mathbb{G}(B)$ (for details on structured perturbations we refer the reader to [15, 16, 17]). However, no general Jordan-like canonical form is known for B-normal matrices, so a different technique to prove this result is inevitable. To present a complete and self-contained treatment for all four matrix classes in the same fashion, proofs of the density results for $\mathbb{J}(B)$, $\mathbb{L}(B)$ and $\mathbb{G}(C)$ are included and presented in the same style based on symmetric polynomials and 1-regularity.

Here we will work with the 2-norm of matrices[1], i.e.

$$\|A\|_2 := \max_{\|x\|_2=1} \|Ax\|_2.$$

The density result stated above is thus equivalent to the fact that, for any $A \in \mathcal{N}(B)$ and any $\varepsilon > 0$, there exists some diagonalizable $A' \in \mathcal{N}(B)$ such that $\|A - A'\|_2 < \varepsilon$. Our motivation for the analysis of density stems from the fact that, for certain B such as

(1.2) $$B = R_n = \begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix} \in \mathbf{M}_n(\mathbb{R}) \quad \text{or} \quad B = J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbf{M}_{2n}(\mathbb{R}),$$

semisimple matrices in $\mathcal{N}(B)$ can all be transformed to sparse and nicely structured (canonical) forms by a B-unitary similarity transformation as shown in [2, Theorems 5 and 8]. From this point of view, our results imply that, even if the transformation established in [2] does not exist for some specific matrix $A \in \mathcal{N}(B)$, it will exist for B-normal matrices arbitrarily close to A.

In Section 2, some preliminaries and basic results are given. Section 3 gives the proof for the main result on the density of semisimple matrices in $\mathcal{N}(B)$ (Section 3.3). The density for the sets $\mathbb{J}(B), \mathbb{L}(B)$ and $\mathbb{G}(B)$ is considered in Sections 3.1 and 3.2. We focus on the case $B = B^H$ in Sections 3.1, 3.2 and 3.3 and discuss the case $B = -B^H$ in Section 3.4. Some conclusions are given in Section 4.

2. Preliminaries and basic results. The set of all $j \times k$ matrices over $\mathbb{F} = \mathbb{R}, \mathbb{C}$ is denoted by $\mathbf{M}_{j \times k}(\mathbb{F})$. Whenever $j = k$, we use the short-hand-notation $\mathbf{M}_k(\mathbb{F}) = \mathbf{M}_{k \times k}(\mathbb{F})$. The notation $\mathbf{Gl}_k(\mathbb{F})$ refers to the general linear group over $\mathbb{F}^k$ (i.e. the set of $k \times k$ nonsingular matrices over $\mathbb{F}$). The set of all eigenvalues of a matrix $A \in \mathbf{M}_k(\mathbb{C})$ is called the spectrum of A and is denoted $\sigma(A)$. The multiplicity of an eigenvalue $\lambda \in \mathbb{C}$ as a root of $\det(xI_k - A)$ is called its algebraic multiplicity and is denoted by $\mathfrak{m}(A, \lambda)$. For each $\lambda \in \sigma(A)$ any vector $v \in \mathbb{C}^k$ that satisfies $Av = \lambda v$ is called an eigenvector for λ. The set of all those vectors corresponding to $\lambda \in \sigma(A)$ is called the eigenspace for λ. It is a subspace of $\mathbb{C}^k$ and its dimension is called the geometric multiplicity of λ. The conjugate transpose of a vector/matrix is denoted with the superscript ^{H} while ^{T} is used for the transposition without complex conjugation. A matrix $A \in \mathbf{M}_n(\mathbb{C})$ is called semisimple (or diagonalizable), if there exists some $S \in \mathbf{Gl}_n(\mathbb{C})$ such that $S^{-1}AS$ is a diagonal matrix.

Let $B \in \mathbf{Gl}_n(\mathbb{C})$. For any $A \in \mathbf{M}_n(\mathbb{C})$ let $A^\star := B^{-1}A^H B$. The matrix $A^\star$ is usually referred to as the adjoint for A, cf. [13, Section 2]. The sets of all B-selfadjoint, B-skewadjoint, B-unitary and B-normal matrices as introduced in Section 1 can now be characterized by the equations $A = A^\star$, $A = -A^\star$, $A^\star = A^{-1}$ and $AA^\star = A^\star A$, respectively. Notice that the mapping $\star : A \mapsto A^\star$ is $\mathbb{R}$-linear, that is, for any $A, C \in \mathbf{M}_n(\mathbb{C})$ and any $\alpha, \gamma \in \mathbb{R}$ it holds that

$$(\alpha A + \gamma C)^\star = \alpha A^\star + \gamma C^\star.$$

---

[1] Our results could be derived in the same way for any submultiplicative and unitarily invariant matrix norm.

Moreover, $(AC)^\star = C^\star A^\star$ holds, so $\star$ is antihomomorphic. The following results on matrices in $\mathbb{J}(B), \mathbb{L}(B)$, $\mathbb{G}(B)$ and $\mathcal{N}(B)$ are well-known and easily established.

REMARK 2.1. Let $B \in \mathbf{Gl}_n(\mathbb{C})$.

(a) The sets $\mathbb{J}(B)$ and $\mathbb{L}(B)$ are $\mathbb{R}$-subspaces of $\mathbf{M}_n(\mathbb{C})$. These sets are often referred to as the Jordan and Lie algebras (cf. [13, Section 2]), respectively, for $[\cdot, \cdot]_B$ since $\mathbb{J}(B)$ is closed under the operation $A \odot C := \frac{1}{2}(AC + CA)$ whereas $\mathbb{L}(B)$ is closed under the Lie bracket $[A, C] := AC - CA$. Moreover, $\mathbb{G}(B)$ is a subgroup of $\mathbf{Gl}_n(\mathbb{C})$ which is called the Lie group.

(b) The sets $\mathbb{J}(B)$ and $\mathbb{L}(B)$ can never be $\mathbb{C}$-subspaces of $\mathbf{M}_n(\mathbb{C})$ since we have, on the one hand, $iJ \in \mathbb{L}(B)$ for any $J \in \mathbb{J}(B)$ and, on the other hand, $iL \in \mathbb{J}(B)$ for any $L \in \mathbb{L}(B)$.

(c) Let $T \in \mathbf{Gl}_n(\mathbb{C})$, $A \in \mathbf{M}_n(\mathbb{C})$ and consider

$$A' := T^{-1}AT \quad \text{and} \quad B' := T^H BT.$$

Then A is B-selfadjoint/B-skewadjoint/B-unitary/B-normal if and only if A' is B'-selfadjoint/B'-skew-adjoint/B'-unitary/B'-normal.

Let $B \in \mathbf{Gl}_n(\mathbb{C})$. Then any $A \in \mathbf{M}_n(\mathbb{C})$ can always be expressed as $A = S + K$ with

$$(2.3) \qquad S := \frac{1}{2}\left(A + A^\star\right) \in \mathbf{M}_n(\mathbb{C}), \qquad K := \frac{1}{2}\left(A - A^\star\right) \in \mathbf{M}_n(\mathbb{C}).$$

Assuming that $B = \pm B^H$, it is easy to see that $S \in \mathbb{J}(B)$ and $K \in \mathbb{L}(B)$. In this case, $A = S + K$ can be interpreted as a B-analogue of the Toeplitz decomposition stated in [9, Theorem 4.1.2]. The next Lemma 2.2 shows that, in case B is additionally unitary (i.e. $B^H B = I_n$ holds), S and K in (2.3) are the best approximations to A from $\mathbb{J}(B)$ and $\mathbb{L}(B)$ with respect to any unitarily invariant matrix norm. The proof of Lemma 2.2 is similar to [3, Theorem 2] (see also [19, Lemma 8.3]).

LEMMA 2.2. Let $B = \pm B^H \in \mathbf{Gl}_n(\mathbb{C})$ be unitary and let $\|\cdot\|$ be any unitarily invariant matrix norm.

(a) Let $A \in \mathbf{M}_n(\mathbb{C})$ and $S := \frac{1}{2}(A + A^\star)$. Then $S \in \mathbb{J}(B)$ and it holds that $\|A - S\| \leq \|A - C\|$ for any other matrix $C \in \mathbb{J}(B)$.

(b) Let $A \in \mathbf{M}_n(\mathbb{C})$ and $K := \frac{1}{2}(A - A^\star)$. Then $K \in \mathbb{L}(B)$ and it holds that $\|A - K\| \leq \|A - C\|$ for any other matrix $C \in \mathbb{L}(B)$.

Proof. (a) Let $\|\cdot\|$ be some unitarily invariant matrix norm and let $A \in \mathbf{M}_n(\mathbb{C})$ be given. It follows from $B = \pm B^H$ that $(A^\star)^\star = A$. Therefore, we have $S^\star = \frac{1}{2}(A + A^\star)^\star = \frac{1}{2}(A^\star + A) = S$, so $S \in \mathbb{J}(B)$. Now let $C \in \mathbb{J}(B)$ be arbitrary. Then

$$\|A - S\| = (1/2)\|A - A^\star\| = \|(A - C) - (A^\star - C)\| = (1/2)\|(A - C) - (A - C)^\star\|$$
$$\leq (1/2)\big(\|A - C\| + \|(A - C)^\star\|\big)$$
$$= (1/2)\big(\|A - C\| + \|(A - C)^H\|\big) = \|A - C\|.$$

In conclusion, we have $\|A - S\| \leq \|A - C\|$. The proof for (b) proceeds along the same lines. $\square$

Let

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \in \mathbf{M}_k(\mathbb{C}),$$

denote a basic Jordan block for λ if $\lambda \in \mathbb{R}$ and define $\hat{J}_k(\lambda) := J_{\tilde{k}}(\lambda) \oplus J_{\tilde{k}}(\overline{\lambda})$ with $2\tilde{k} = k$ in case $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Here and in the following, the notation $\oplus$ is used to denote the direct sum of two matrices, i.e. $X \oplus Y = \text{diag}(X, Y)$. The next two theorems for the case when $B = B^H$ are taken from [4] and will be useful for some proofs in Section 3. For convenience, we omit the subscript k for the Jordan blocks in Theorem 2.3.

THEOREM 2.3 ([4, Theorem 5.1.1]).   Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$ and $A \in \mathbb{J}(B)$.   Then there is some $T \in \mathbf{Gl}_n(\mathbb{C})$ such that $T^{-1}AT = J$ and $T^H BT = \widetilde{B}$ where

$$J = J(\lambda_1) \oplus \cdots \oplus J(\lambda_\alpha) \oplus \hat{J}(\lambda_{\alpha+1}) \oplus \cdots \oplus \hat{J}(\lambda_\beta),$$

is a Jordan normal form for A where $\lambda_1, \ldots, \lambda_\alpha$ are the real eigenvalues of A and $\lambda_{\alpha+1}, \ldots, \lambda_\beta$ are the nonreal eigenvalues of A from the upper half-plane. Moreover,

$$\widetilde{B} := T^H BT = \eta_1 P_1 \oplus \cdots \oplus \eta_\alpha P_\alpha \oplus P_{\alpha+1} \oplus \cdots \oplus P_\beta,$$

where $P_k = R_p$ is the $p \times p$ reverse identity matrix (see also (1.2)) with $p \times p$ being the size of $J(\lambda_k)$ $(k = 1, \ldots, \alpha)$ or $\hat{J}(\lambda_k)$ $(k = \alpha + 1, \ldots, \beta)$ and $\eta = \{\eta_1, \ldots, \eta_\alpha\}$ is an ordered set of signs ± 1.

The transformation defined in Theorem 2.4 can be interpreted as a B-analogue of the well-known Cayley transform [5]. It can be used to turn B-selfadjoint matrices into B-unitary matrices and vice versa.

THEOREM 2.4 ([4, Proposition 4.3.4]).   Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$ and $A \in \mathbb{J}(B)$. Let $w \in \mathbb{C}$ be a nonreal number with $w \notin \sigma(A)$ and let $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Then

$$(2.4) \qquad\qquad U = \alpha(A - \overline{w}I_n)(A - wI_n)^{-1} \in \mathbb{G}(B),$$

and $\alpha \notin \sigma(U)$. Conversely, if $U \in \mathbb{G}(B)$, $|\alpha| = 1$ and $\alpha \notin \sigma(U)$, then for any $w \neq \overline{w}$ we have

$$(2.5) \qquad\qquad A = (wU - \overline{w}\alpha I_n)(U - \alpha I_n)^{-1} \in \mathbb{J}(B),$$

and $w \notin \sigma(A)$. The formulas (2.4) and (2.5) are inverse to one other.

2.1. Symmetric polynomials and 1-regularity. In this section, some mathematical tools and basic results are presented that are used for the proofs in Section 3. A matrix $A \in \mathbf{M}_n(\mathbb{C})$ is called 1-regular[2], if the eigenspace for any $\lambda \in \sigma(A)$ is one-dimensional. In particular, A is 1-regular if and only if there is only one basic Jordan block for each eigenvalue λ in the Jordan decomposition of A. This implies that a diagonalizable matrix is 1-regular if and only if all its eigenvalues are (pairwise) distinct. Moreover, it follows directly that 1-regularity is preserved under similarity transformations. The following Theorem 2.5 characterizes 1-regular matrices and is of central importance in the further discussion.

THEOREM 2.5 ([18, Proposition 1.1.2]).   Let $A \in \mathbf{M}_n(\mathbb{C})$.   Then each of the following statements is equivalent to A being 1-regular.

(a) The centralizer of A in $\mathbf{M}_n(\mathbb{C})$ coincides with $\mathbb{C}[A]$.   That is, for all $A' \in \mathbf{M}_n(\mathbb{C})$ commuting with A, there is some polynomial $p(x) \in \mathbb{C}[x]$ such that $A' = p(A)$.
(b) The dimension of $\mathbb{C}[A]$ is n, i.e. the matrices $I_n, A, A^2, \ldots, A^{n-1}$ are linearly independent over $\mathbb{C}$.

The next Proposition 2.6 shows that, for any given $A \in \mathbb{J}(B)$, we can always find a 1-regular matrix from $\mathbb{J}(B)$ that commutes with A. From now one we confine ourselves to the case $B = B^H$ and postpone the discussion of the situation for $B = -B^H$ to Subsection 3.4.

---

[2]A 1-regular matrix is sometimes also called nonderogatory.

PROPOSITION 2.6. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$ and $A \in \mathbb{J}(B)$. Then there exists a 1-regular matrix $C \in \mathbb{J}(B)$ such that A and C commute.*

Proof. We apply Theorem 2.3 to A and B (the subscript indicating the size of the Jordan blocks is omitted). So, there exists some $T \in \mathbf{Gl}_n(\mathbb{C})$ such that

$$\widetilde{A} := T^{-1}AT = J(\lambda_1) \oplus \cdots \oplus J(\lambda_\alpha) \oplus \hat{J}(\lambda_{\alpha+1}) \oplus \cdots \oplus \hat{J}(\lambda_\beta),$$

where $\lambda_1, \ldots, \lambda_\alpha \in \mathbb{R}$ are the real eigenvalues of A and $\lambda_{\alpha+1}, \ldots, \lambda_\beta \in \mathbb{C}$ are the nonreal eigenvalues of A from the upper half-plane. Moreover,

$$\widetilde{B} := T^H BT = \eta_1 P_1 \oplus \cdots \eta_\alpha P_\alpha \oplus P_{\alpha+1} \oplus \cdots \oplus P_\beta,$$

where $P_k = R_p$ is the $p \times p$ reverse identity matrix with $p \times p$ being the size of $J(\lambda_k)$ ($k = 1, \ldots, \alpha$) or $\hat{J}(\lambda_k)$ ($k = \alpha + 1, \ldots, \beta$), and each η_j is either $+1$ or -1. According to Remark 2.1(c), we have $\widetilde{A} \in \mathbb{J}(\widetilde{B})$ since $A \in \mathbb{J}(B)$. Now let $a_1, \ldots, a_\alpha \in \mathbb{R}$ and $a_{\alpha+1}, \ldots, a_\beta \in \mathbb{C} \setminus \mathbb{R}$ be arbitrary and pairwise distinct values and consider

$$\widetilde{C} := J(a_1) \oplus \cdots \oplus J(a_\alpha) \oplus \hat{J}(a_{\alpha+1}) \oplus \cdots \oplus \hat{J}(a_\beta),$$

where each $J(a_k)$ has the same size as $J(\lambda_k)$, $k = 1, \ldots, \alpha$, and each $\hat{J}(a_k)$ has the same size as $\hat{J}(\lambda_k)$, $k = \alpha + 1, \ldots, \beta$. Observe that $\widetilde{C} \in \mathbb{J}(\widetilde{B})$ and that $\widetilde{A}\widetilde{C} = \widetilde{C}\widetilde{A}$ holds. Moreover, $\widetilde{C}$ is 1-regular since the values a_k are all distinct. We now apply the reverse transformation to obtain

$$A = T\widetilde{A}T^{-1}, \quad C := T\widetilde{C}T^{-1}, \quad B = T^{-H}\widetilde{B}T^{-1}.$$

Note that now $AC = CA$ holds and that $C \in \mathbb{J}(B)$ (according to Remark 2.1(c) since we had $\widetilde{C} \in \mathbb{J}(\widetilde{B})$). Finally, as $\widetilde{C}$ was 1-regular so is C and the proof is complete. □

Using Remark 2.1(b), the result from Proposition 2.6 can easily be extended to $\mathbb{L}(B)$.

COROLLARY 2.7. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$ and $A \in \mathbb{L}(B)$. Then there exists a 1-regular matrix $C \in \mathbb{L}(B)$ such that A and C commute.*

Proof. This follows immediately from Remark 2.1(b) and Proposition 2.6. Additionally, notice that $C \in \mathbf{M}_n(\mathbb{C})$ is 1-regular if and only if iC is 1-regular. □

A polynomial $p(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n]$ in $n \geq 1$ unknowns is called symmetric if

$$p(x_1, \ldots, x_n) = p(x_{\tau(1)}, \ldots, x_{\tau(n)}),$$

holds for all permutations τ of $1, 2, \ldots, n$. The following Theorem 2.8 will be central in the next section.

THEOREM 2.8 ([18, Proposition 7.1.10]). *Let $p(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n]$ be a symmetric polynomial and $f : \mathbf{M}_n(\mathbb{C}) \to \mathbb{C}$ be a function given by*

$$f(A) = p\big(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)\big) =: p(A),$$

where $\lambda_k(A), k = 1, \ldots, n$, denote the eigenvalues of A. Then there is a polynomial $q(x_{11}, x_{12}, \ldots, x_{nn}) \in \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}]$ in n^2 unknowns such that

$$f(A) = q\big(a_{11}, a_{12}, \ldots, a_{nn}\big) \quad \forall\, A = [a_{i,j}]_{i,j} \in \mathbf{M}_n(\mathbb{C}).$$

Proof. The proof follows the one from [18]. Let $A = [a_{i,j}]_{i,j} \in \mathbf{M}_n(\mathbb{C})$ and let $\chi_A(x) = \det(xI_n - A)$ be the characteristic polynomial of A. Assume that

$$\chi_A(x) = x^n + c_1 x^{n-1} + \cdots + c_{n-1}x + c_n, \tag{2.6}$$

for some $c_k \in \mathbb{C}, k = 1, \ldots, n$. It is well known that $c_1, \ldots, c_n$ are polynomials in the entries $a_{ij}, 1 \le i, j \le n$, of A. Furthermore, let $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ denote the eigenvalues of A. Then, $\chi_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ and its expansion gives

$$\chi_A(x) = x^n - s_1 x^{n-1} + s_2 x^{n-2} + \cdots + (-1)^n s_n, \tag{2.7}$$

for some coefficients $s_1, \ldots, s_n \in \mathbb{C}$. A closer inspection reveals that $s_1, \ldots, s_n$ are given by the n elementary symmetric polynomials in $\lambda_1, \ldots, \lambda_n$ (see also [10, Section 2]), that is

$$(2.8) \quad s_1 = \sum_i \lambda_i \ (= \operatorname{trace}(A)), \qquad s_2 = \sum_{i<j} \lambda_i \lambda_j, \qquad s_3 = \sum_{i<j<k} \lambda_i \lambda_j \lambda_k, \quad \ldots \quad s_n = \prod_i \lambda_i \ (= \det(A)).$$

A comparison of the coefficients in the expressions (2.6) and (2.7) yields: each elementary symmetric polynomial $s_k = s_k(\lambda_1, \ldots, \lambda_n)$ in (2.8) in the eigenvalues of A agrees with a certain polynomial $c_k = p_k(a_{11}, a_{12}, \ldots, a_{nn})$ in the entries a_{ij} of A. In consequence, this is true for any symmetric polynomial $q(\lambda_1, \ldots, \lambda_n)$ since each such q can always be expressed as a polynomial in $s_1, \ldots, s_n$ (cf. [10, Theorem 2.20]). This completes the proof. $\square$

EXAMPLE 2.9. We give three applications of Theorem 2.8 that will be important in the further discussion. Each of these examples can be found in [18, Section 7]. First, let $A = [a_{ij}]_{ij} \in \mathbf{M}_n(\mathbb{C})$ with eigenvalues $\lambda_1(A), \ldots, \lambda_n(A)$.

(a) The function $p : \mathbf{M}_n(\mathbb{C}) \to \mathbb{C}$ given by

$$p(A) := p(\lambda_1(A), \ldots, \lambda_n(A)) = \prod_k \lambda_k(A),$$

determines whether A is invertible. That is, $p(A) = 0$ if A is singular and $p(A) \ne 0$ otherwise. As p is a symmetric polynomial in $\lambda_1(A), \ldots, \lambda_n(A)$, there is some $q(x_{11}, \ldots, x_{nn})$ is n^2 unknowns with $q(A) := q(a_{11}, a_{12}, \ldots, a_{nn}) = p(A)$ for all $A \in \mathbf{M}_n(\mathbb{C})$. The polynomial q in $a_{11}, a_{12}, \ldots, a_{nn}$ is given by the determinant.

(b) The function $p : \mathbf{M}_n(\mathbb{C}) \to \mathbb{C}$ given by

$$p(A) := p(\lambda_1(A), \ldots, \lambda_n(A)) = \prod_{k \ne j} (\lambda_k(A) - \lambda_j(A)),$$

determines whether A has a multiple eigenvalue. That is, $p(A) \ne 0$ if all eigenvalues of A are (pairwise) distinct and $p(A) = 0$ otherwise. Notice that $p(\lambda_1(A), \ldots, \lambda_n(A))$ is a symmetric polynomial in $\lambda_1(A), \ldots, \lambda_n(A)$. According to Theorem 2.8, there exists a polynomial $q \in \mathbb{C}[x_{11}, \ldots, x_{nn}]$ in n^2 unknowns such that

$$q(A) = q(a_{11}, a_{12}, \ldots, a_{nn}) = p(\lambda_1(A), \ldots, \lambda_n(A)) = p(A),$$

for all $A = [a_{ij}]_{ij} \in \mathbf{M}_n(\mathbb{C})$.

(c) Now assume that $A = [a_{ij}]_{ij} \in \mathbf{M}_{m \times n}(\mathbb{C})$. Recall that a $k \times k$ minor of A is the determinant of a submatrix $A' \in \mathbf{M}_k(\mathbb{C})$ obtained from A by deleting $m - k$ rows and $n - k$ columns. Thus (the value of) each minor is expressible as a polynomial in some entries of A according to (a). Now note that for A the fact $\operatorname{rank}(A) \leq r < \min\{m, n\}$ is equivalent to the vanishing of all $(r + 1) \times (r + 1)$ minors of A. As there are

$$s := \binom{m}{r+1} \cdot \binom{n}{r+1},$$

such minors of A, according to Theorem 2.8, there are s polynomials $q_k(x_{11}, \ldots, x_{mn}), 1 \leq k \leq s$, in mn unknowns $x_{11}, x_{12}, \ldots, x_{mn}$ such that $q_k(A) := q_k(a_{11}, a_{12}, \ldots, a_{mn}) = 0$ holds for all $k = 1, \ldots, s$ if and only if $\operatorname{rank}(A) \leq r$. In particular, if $\operatorname{rank}(A) > r$, there is at least one q_k such that $q_k(A) = q_k(a_{11}, \ldots, a_{mn}) \neq 0$.

Do not overlook that, in each case considered in Example 2.9, the polynomials that have been determined do not depend on a special matrix.

Recall that, according to Theorem 2.5, $A \in \mathbf{M}_n(\mathbb{C})$ is 1-regular if and only if $I_n, A, A^2, \ldots, A^{n-1}$ are linearly independent. In turn, this is the case if and only if the matrix

$$M := \begin{bmatrix} | & | & | & & | \\ I_n & A & A^2 & \cdots & A^{n-1} \\ | & | & | & & | \end{bmatrix} \in \mathbf{M}_{n^2 \times n}(\mathbb{C}),$$

whose columns are $I_n, A, A^2, \ldots$ written in vectorized fashion as $n^2 \times 1$ column vectors, has full rank (i.e. $\operatorname{rank}(M) = n$). Certainly, the entries of M are polynomials in the entries of A. The matrix M has rank $\leq n - 1$ if and only if all $n \times n$ minors of M vanish simultaneously. Taking Example 2.9(c) into account, we have the following result:

COROLLARY 2.10. *There exists a collection of $s \geq 1$ (nonzero) polynomials*

$$w_k(x_{11}, \ldots, x_{nn}) \in \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}], \quad k = 1, \ldots, s,$$

in n^2 unknowns such that $w_k(A) = w_k(a_{11}, \ldots, a_{nn}) = 0$ holds for $A = [a_{ij}]_{ij} \in \mathbf{M}_n(\mathbb{C})$ and all w_k if and only if A is not 1-regular.

In other words, Corollary 2.10 states that $w_\ell(A) \neq 0$ for at least one $\ell, 1 \leq \ell \leq s$, is a sufficient condition for $A \in \mathbf{M}_n(\mathbb{C})$ to be 1-regular. This fact will be very useful for the proofs in the upcoming section.

3. Density of diagonalizable matrices. In this section, we prove our main theorem on the density of diagonalizable matrices in $\mathcal{N}(B)$ (Section 3.3) using the results obtained on symmetric polynomials and 1-regular matrices from Section 2.1. The proof will particularly make use of the density of semisimple matrices for the set $\mathbb{J}(B)$ which is proven first (Section 3.1). Applying this result in combination with the Cayley transform (see Theorem 2.4) directly provides a proof for the set of B-unitary matrices that is considered in Section 3.2. Finally, the case $B = -B^H$ is treated in Section 3.4.

3.1. The Lie and Jordan algebras. We begin by considering the density of diagonalizable matrices in the sets of B-selfadjoint and B-skewadjoint matrices. The result of Proposition 3.1 will be central for the proof of Theorem 3.2.

PROPOSITION 3.1. *Let $B = B^H \in \mathbf{M}_n(\mathbb{C})$ be nonsingular. Then $\mathbb{J}(B)$ and $\mathbb{L}(B)$ contain a matrix with pairwise distinct eigenvalues.*

Proof. Let (m_-, m_+) be the inertia[3] of B, that is, denote the number of positive real eigenvalues of B by m_+ and the number of negative real eigenvalues of B by m_-. According to a Theorem of Sylvester (cf. [9, Theorem 4.5.7]), there exists some $Q \in \mathbf{Gl}_n(\mathbb{C})$ such that

$$B' := Q^H BQ = \begin{bmatrix} -I_{m_-} & \\ & I_{m_+} \end{bmatrix}.$$

Now for any n distinct values $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ the matrix $D = \mathrm{diag}(\alpha_1, \ldots, \alpha_n) \in \mathbf{M}_n(\mathbb{C})$ is in $\mathbb{J}(B')$. Thus, $A := QDQ^{-1}$ is in $\mathbb{J}(Q^{-H}(Q^H BQ)Q^{-1}) = \mathbb{J}(B)$ according to Remark 2.1(c) and has pairwise distinct eigenvalues. The same statement for $\mathbb{L}(B)$ follows by taking the diagonal entries $i\alpha_1, \ldots, i\alpha_n$ for D. □

We now prove the density statement of semisimple matrices in $\mathbb{J}(B)$ and $\mathbb{L}(B)$. The proof makes use of the fact observed in Example 2.9(b). With the use of Proposition 3.1, it can also be derived from [8, Corollary 1].

THEOREM 3.2. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$.*

(a) *For any $J \in \mathbb{J}(B)$ and any $\varepsilon > 0$ there exists some diagonalizable $J' \in \mathbb{J}(B)$ such that $\|J - J'\|_2 < \varepsilon$. Moreover, J' can be chosen to have pairwise distinct eigenvalues.*
(b) *For any $L \in \mathbb{L}(B)$ and any $\varepsilon > 0$ there exists some diagonalizable $L' \in \mathbb{L}(B)$ such that $\|L - L'\|_2 < \varepsilon$. Moreover, L' can be chosen to have pairwise distinct eigenvalues.*

Proof. Let $A = [a_{ij}]_{ij} \in \mathbf{M}_n(\mathbb{C})$. Recall that, according to Example 2.9(b), there is some polynomial $q(x_{11}, \ldots, x_{nn}) \in \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}]$ in n^2 unknowns x_{ij}, $1 \le i, j \le n$, such that $q(A) = q(a_{11}, \ldots, a_{nn}) = 0$ if and only if A has a (i.e. at least one) multiple eigenvalue. Otherwise, that is if all eigenvalues of A are (pairwise) distinct, $q(A) \ne 0$.

(a) Now assume $J \in \mathbb{J}(B)$. According to Proposition 3.1, there is some $E \in \mathbb{J}(B)$ with distinct eigenvalues. Now consider the family of matrices $M(z) := zJ + E$, $z \in \mathbb{C}$, and the polynomial $\widetilde{q}(z) := q(M(z)) = q(zJ + E) \in \mathbb{C}[z]$ that only depends on the single variable z. Certainly, $\widetilde{q}(0) = q(E) \ne 0$ since the eigenvalues of E are distinct. Therefore, $\widetilde{q} \ne 0$ and $\widetilde{q}$ is not the zero-polynomial. As $\widetilde{q}(z)$ has only a finite number of roots, almost all matrices $M(z_0), z_0 \in \mathbb{C}$, have distinct eigenvalues. Consequently, the same holds for all $J + cE = c(c^{-1}J + E)$, $c \in \mathbb{C}$. To guarantee that $J + cE \in \mathbb{J}(B)$, we confine ourselves to the case $c \in \mathbb{R}$ from here on (c.f. Remark 2.1(b)). Next, let $\varepsilon > 0$ be given and choose some $c_0 \in \mathbb{R}$ with $|c_0| < \varepsilon/\|E\|_2$ such that $\widetilde{q}(c_0^{-1}) \ne 0$. Then $J' := J + c_0 E \in \mathbb{J}(B)$, J' has n distinct eigenvalues and

$$\|J - J'\|_2 = |c_0| \cdot \|E\|_2 < \frac{\varepsilon}{\|E\|_2} \cdot \|E\|_2 = \varepsilon.$$

Moreover, as J' has pairwise distinct eigenvalues it is diagonalizable.

(b) The proof for $L \in \mathbb{L}(B)$ proceeds along the same lines. □

COROLLARY 3.3. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. The set of matrices with pairwise distinct eigenvalues is dense in $\mathbb{J}(B)$ and $\mathbb{L}(B)$.*

---

[3] As $B \in \mathbf{Gl}_n(\mathbb{C})$, we have $0 \notin \sigma(B)$. Moreover, as $B = B^H$, all eigenvalues of B are real. In conclusion, B has only positive and negative real eigenvalues.

REMARK 3.4. As mentioned in Section 1, the statement of Corollary 3.3 regarding $\mathbb{J}(B)$ can also be derived from the canonical in Theorem 2.3. In particular, assume that $T^{-1}AT = J$ and $T^H BT = \widetilde{B}$ where

$$J = J(\lambda_1) \oplus \cdots \oplus J(\lambda_\alpha) \oplus \hat{J}(\lambda_{\alpha+1}) \oplus \cdots \oplus \hat{J}(\lambda_\beta),$$

as in Theorem 2.3. Adding some $\varepsilon > 0$ in the lower left corner of a Jordan block $J_k(\lambda)$ gives

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ \varepsilon & & & & \lambda \end{bmatrix} \in \mathbf{M}_k(\mathbb{C}),$$

which is a matrix whose eigenvalues are all distinct and lie on a circle with radius $\varepsilon^{1/k}$ around λ. Proceeding in this fashion for appropriately chosen small values of ε for each Jordan block, one obtains a semisimple matrix $\tilde{J}$ that is still in $\mathbb{J}(B)$. The appropriate choice of the perturbation yields that $\|J - \tilde{J}\|_2$ can in fact be made arbitrarily small and the statement of Theorem 3.2(a) follows. The result for $\mathbb{L}(B)$ can also be derived similarly.

The proof of Theorem 3.2 also reveals that any matrix from $\mathbb{J}(B)$ or $\mathbb{L}(B)$ can always be expressed as a sum of matrices with pairwise distinct eigenvalues from the same class.

COROLLARY 3.5. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. Then every matrix in $\mathbb{J}(B)$ can be expressed as a sum of two diagonalizable matrices from $\mathbb{J}(B)$ with pairwise distinct eigenvalues. The same holds for $\mathbb{L}(B)$.*

Proof. Using the notation from the proof of Theorem 3.2(a), $E \in \mathbb{J}(B)$ and $c_0 \in \mathbb{R}$ can be chosen such that E and $J + c_0 E$ both have n distinct eigenvalues. Then, $J = (J + c_0 E) - c_0 E$ is a sum of two matrices $J + c_0 E, -c_0 E \in \mathbb{J}(B)$ which both have pairwise distinct eigenvalues. The proof follows analogously for $\mathbb{L}(B)$. ☐

Let $A \in \mathbf{M}_n(\mathbb{C})$. Using the decomposition $A = S + K$ with $S = (1/2)(A + A^\star)$ and $K = (1/2)(A - A^\star)$, see (2.3), accompanied by Corollary 3.5 we end this section with the following observation related to the set $\mathcal{N}(B)$ of B-normal matrices.

PROPOSITION 3.6. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. Any matrix $A \in \mathbf{M}_n(\mathbb{C})$ can be expressed as a sum of four matrices from $\mathcal{N}(B)$ with each having pairwise distinct eigenvalues.*

Proof. Any matrix from $\mathbb{J}(B)$ and $\mathbb{L}(B)$ can be expressed as a sum of two matrices with n distinct eigenvalues from the same class, cf. Corollary 3.5. This result can directly be applied to $A = S + K$ with $S = (1/2)(A + A^\star) \in \mathbb{J}(B)$ and $K = (1/2)(A - A^\star) \in \mathbb{L}(B)$. ☐

3.2. The set $\mathbb{G}(B)$ of B-unitary matrices. In this section, we analyze the set $\mathbb{G}(B)$ and show that it contains a dense subset of diagonalizable matrices. On the one hand, a perturbation-based proof can be formulated using the structured canonical Jordan-like form for B-unitary matrices, see [4, Section 5] or [6], following the same ideas as in Remark 3.4. On the other hand, the result from Theorem 3.2(a) can be used along with the Cayley transformation (cf. Theorem 2.4) to prove the result.

THEOREM 3.7. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. For any $G \in \mathbb{G}(B)$ and any $\varepsilon > 0$ there exists some diagonalizable $G' \in \mathbb{G}(B)$ such that $\|G - G'\|_2 \leq \varepsilon$.*

Proof. Assume $G \in \mathbb{G}(B)$ is given. Let $\alpha \in \mathbb{C}$, $|\alpha| = 1$, be chosen such that $\alpha \notin \sigma(G)$ and let $w \neq \overline{w}$ be some fixed number. Then, according to Theorem 2.4,

$$G' := \left(wG - \overline{w}\alpha I_n\right)\left(G - \alpha I_n\right)^{-1} \in \mathbb{J}(B),$$

with $w \notin \sigma(G')$. By Theorem 3.2 we may construct a sequence $(F'_k)_k \in \mathbb{J}(B)$ of diagonalizable matrices with $F'_k \to G'$ for $k \to \infty$. Again according to Theorem 2.4, $F_k := \alpha(F'_k - \overline{w}I_n)(F'_k - wI_n)^{-1} \to G$ for $k \to \infty$ since the transformation is continuous (and both transformations in Theorem 2.4 are inverse to each other). Certainly, $w \notin \sigma(F'_k)$ has to hold for F_k to be defined. However, $w \notin \sigma(G')$ implies that $w \notin \sigma(F'_k)$ will hold if F'_k is close enough to G' (this can be interpreted as a consequence of the Bauer-Fike Theorem, cf. [5, Theorem 7.2.2], since all F'_k are diagonalizable). Formally, there is some $\eta > 0$ such that $w \notin \sigma(F'_k)$ for all $F'_k \in \mathbf{M}_n(\mathbb{C})$ with $\|G' - F'_k\|_2 < \eta$. From now on, it suffices to consider only those F'_k from the sequence which are close enough to G' such that F_k is defined. Now let $\varepsilon > 0$ be given. Then there exists some $\delta > 0$ such that

$$\|F_k - G\|_2 < \varepsilon \quad \text{for all } F_k \text{ such that} \quad \|F'_k - G'\|_2 < \delta,$$

due to the continuity of the transformation. Now choose some F'_j from the sequence with $\|F'_j - G'\|_2 < \min\{\delta, \eta\}$ (so, in particular, $w \notin \sigma(F'_j)$). Then, $F_j \in \mathbb{G}(B)$ is defined and $\|F_j - G\|_2 < \varepsilon$. As F'_j is diagonalizable, assume $S^{-1}F'_j S = D$ for some diagonal $D \in \mathbf{M}_n(\mathbb{C})$. Then it follows from a direct calculation that

$$(3.9) \qquad\qquad S^{-1}F_j S = \alpha\left(D - \overline{w}I_n\right)\left(D - wI_n\right)^{-1},$$

is a diagonalization of F_j and the proof is complete. $\square$

COROLLARY 3.8. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. The set of matrices with pairwise distinct eigenvalues is dense in $\mathbb{G}(B)$.*

Proof. This follows from the proof of Theorem 3.7 and Corollary 3.3, since the sequence of matrices $(F'_k)_k \in \mathbb{J}(B)$ constructed in the proof of Theorem 3.7 can be chosen such that all matrices have pairwise distinct eigenvalues. Then, if $S^{-1}F'_j S = D$ for some diagonal $D \in \mathbf{M}_n(\mathbb{C})$ with n distinct eigenvalues, the matrix $S^{-1}F_j S$ in (3.9) has distinct eigenvalues, too. $\square$

3.3. The set $\mathcal{N}(B)$ of B-normal matrices. Now we finally consider the proof for the density of diagonalizable matrices in the set $\mathcal{N}(B)$ of B-normal matrices. The following Lemma 3.9 will be helpful to prove our main Theorem 3.10. It shows how to construct B-normal matrices from pairs of commuting B-selfadjoint matrices.

LEMMA 3.9. *Let $B = B^H \in \mathbf{Gl}_n(\mathbb{C})$. If $F, G \in \mathbb{J}(B)$ and F and G commute, then $A = F \pm iG \in \mathcal{N}(B)$.*

Proof. Note that

$$A^\star = (F \pm iG)^\star = B^{-1}(F \pm iG)^H B = B^{-1}F^H B \pm B^{-1}(iG)^H B = F \mp iG,$$

so $AA^\star = (F \pm iG)(F \mp iG)$ and $A^\star A = (F \mp iG)(F \pm iG)$. Since $FG = GF$ we see that $AA^\star = A^\star A$ holds. $\square$

Theorem 3.10 states that the density result obtained for $\mathbb{J}(B), \mathbb{L}(B)$ and $\mathbb{G}(B)$ before is true for $\mathcal{N}(B)$ under the additional assumption that B is a unitary matrix. Once Theorem 3.10 is proven, we use it to prove the general case without the unitarity assumption.

THEOREM 3.10. *Let $B \in \mathbf{Gl}_n(\mathbb{C})$ with $B = B^H$ and $B^H B = I_n$. For any $N \in \mathcal{N}(\mathrm{B})$ and any $\varepsilon > 0$ there exists some diagonalizable $N' \in \mathcal{N}(\mathrm{B})$ such that $\|N - N'\|_2 \leq \varepsilon$.*

Proof. Let $N \in \mathcal{N}(B)$ be arbitrary. We define $S := \frac{1}{2}(N + N^\star) \in \mathbb{J}(B)$, $K := \frac{1}{2}(N - N^\star) \in \mathbb{L}(B)$ as in (2.3) and express N as

$$(3.10) \qquad N = S + K = S - i^2 K = S - iK_H,$$

with $K_H := iK$. Notice that $K_H \in \mathbb{J}(B)$ according to Remark 2.1(b). It follows straight forward that $SK_H = K_H S$ holds, that is

$$\begin{aligned} SK_H &= (1/2)\big(A + A^\star\big) \cdot (i/2)\big(A - A^\star\big) = (i/4)\big(A + A^\star\big)\big(A - A^\star\big) \\ &= (i/4)\big(A - A^\star\big)\big(A + A^\star\big) = (i/2)\big(A - A^\star\big) \cdot (1/2)\big(A + A^\star\big) \\ &= K_H S, \end{aligned}$$

so S and K_H commute. According to Proposition 2.6, there exists some 1-regular $E = [e_{ij}]_{ij} \in \mathbb{J}(B)$ such that K_H and E commute, that is, $K_H E = E K_H$ holds. Now, we consider the family of all matrices $M = M(z) = zS + E \in \mathbf{M}_n(\mathbb{C})$ for $z \in \mathbb{C}$. As both S and E commute with K_H, so does each $M(z)$. In particular, note that $M(0) = E$ is 1-regular. Now, according to Corollary 2.10, there is some $w_\ell(x_{11}, \ldots, x_{nn})$ (from the set of polynomials w_k, $k = 1, \ldots, s$, that vanish simultaneously for matrices that are not 1-regular) such that $w_\ell(E) := w_\ell(e_{11}, e_{12}, \ldots, e_{nn}) \neq 0$. Now, as S and E are fixed, consider $\widetilde{w}(z) := w_\ell(M(z)) = w_\ell(zS + E) \in \mathbb{C}[z]$ as a single-variable-polynomial and notice that $\widetilde{w}(0) = w_\ell(E) \neq 0$. Thus, $\widetilde{w} \neq 0$ is not the zero-polynomial. Recall that $\widetilde{w}(z_0) \neq 0$ is a sufficient condition for $M(z_0) = z_0 S + E$ to be 1-regular. Consequently, as $\widetilde{w}(z)$ does only have a finite number of roots, $M(z_0) = z_0 S + E$ will be 1-regular for almost all $z_0 \in \mathbb{C}$. Therefore, $S + cE = c(c^{-1}S + E)$ is also 1-regular for all but a finite number of nonzero $c \in \mathbb{C}$.

Now, let $\varepsilon = 2\epsilon > 0$ be given. Choose some $c \in \mathbb{R}$ with $|c| \leq \epsilon/(2\|E\|_2)$ such that $S_c := S + cE \in \mathbb{J}(B)$ is 1-regular. Then,

$$\|S - S_c\|_2 = \|cE\|_2 = |c|\|E\|_2 \leq \frac{\epsilon}{2\|E\|_2}\|E\|_2 = \frac{\epsilon}{2}.$$

As S_c and K_H commute (recall that S and E both commute with K_H) and S_c is 1-regular, there exists some polynomial $p(x) \in \mathbb{C}[x]$ with $p(S_c) = K_H$ according to Theorem 2.5(a). Moreover, from Theorem 3.2 (see also Corollary 3.3), there exists a sequence $(F_k)_k \in \mathbb{J}(B)$, $k \in \mathbb{N}$, of diagonalizable matrices with $F_k \to S_c$. Thus, $p(F_k) \to K_H$ for $k \to \infty$ since $p(x)$ is continuous. Now, for $\epsilon/2$ there exists some $\delta > 0$ such that

$$(3.11) \qquad \|p(S_c) - p(F_k)\|_2 = \|K_H - p(F_k)\|_2 < \frac{\epsilon}{2} \quad \text{if} \quad \|S_c - F_k\|_2 < \delta,$$

due to the continuity of $p(x)$. Next, choose some $F_\ell \in \mathbb{J}(B)$ from the sequence $(F_k)_k$ with $\|S_c - F_\ell\| < \min\{\epsilon/2, \delta\}$. Then, we obtain that

$$(3.12) \qquad \begin{aligned} \|S - F_\ell\|_2 &= \|S - S_c + S_c - F_\ell\|_2 \\ &\leq \|S - S_c\|_2 + \|S_c - F_\ell\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

As $p(x) \in \mathbb{C}[x]$ may have complex coefficients, it might be the case that $p(F_\ell) \notin \mathbb{J}(B)$. However, $G := (1/2)(p(F_\ell) + p(F_\ell)^\star) \in \mathbb{J}(B)$ (recall (2.3)). Assume that $p(x)$ is given by $\sum_{k=0}^t a_k x^k$ for complex coefficients $a_k \in \mathbb{C}, k = 0, \ldots, t$. Then,

$$p(F_\ell)^\star = B^{-1}\left(\sum_{k=0}^t a_k F_\ell^k\right)^H B = B^{-1}\left(\sum_{k=0}^t \overline{a_k}\left(F_\ell^k\right)^H\right)B$$

$$= B^{-1}\left(\sum_{k=0}^t \overline{a_k}\left(F_\ell^H\right)^k\right)B = \sum_{k=0}^t \overline{a_k}\left(B^{-1}F_\ell^H B\right)^k = \sum_{k=0}^t \overline{a_k}F_\ell^k,$$

so $p(F_\ell)^\star = q(F_\ell)$ with $q(x) = \sum_{k=0}^t \overline{a_k}x^k$. In particular, $(1/2)(p(F_\ell) + q(F_\ell)) = r(F_\ell) = \sum_{k=0}^t 2\mathfrak{Re}(a_k)F_\ell^k$ is a (real) polynomial in F_ℓ (where $\mathfrak{Re}(a_k)$ denotes the real part of the complex number a_k). As

$$\|p(F_\ell) - \frac{1}{2}(p(F_\ell) + p(F_\ell)^\star)\|_2 = \|p(F_\ell) - r(F_\ell)\|_2 \le \|p(F_\ell) - X\|_2,$$

holds for any $X \in \mathbb{J}(B)$ according to Lemma 2.2 and $\|p(F_\ell) - K_H\|_2 < \epsilon/2$ (recall (3.11) and the choice of F_ℓ) we conclude $\|p(F_\ell) - r(F_\ell)\|_2 \le \epsilon/2$. In analogy to (3.12), we obtain

$$\|K_H - G\|_2 = \|K_H - p(F_\ell) + p(F_\ell) - G\|_2$$

$$\le \|K_H - p(F_\ell)\|_2 + \|p(F_\ell) - G\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Finally, we arrived at $\|S - F_\ell\|_2 < \epsilon$ and $\|K_H - r(F_\ell)\|_2 < \epsilon$. As $F_\ell, r(F_\ell) \in \mathbb{J}(B)$, note that $N' := F_\ell - ir(F_\ell) \in \mathcal{N}(B)$ according to Lemma 3.9. Moreover, we have

$$\|N - N'\|_2 = \|(S - iK_H) - (F_\ell - ir(F_\ell))\|_2 = \|(S - F_\ell) - i(K_H - r(F_\ell)\|_2$$

$$\le \|S - F_\ell\|_2 + \|K_H - r(F_\ell)\|_2 \le \epsilon + \epsilon = 2\epsilon = \varepsilon.$$

As F_ℓ is diagonalizable, so is $r(F_\ell)$. In addition, as F_ℓ and $r(F_\ell)$ commute, they are simultaneously diagonalizable [9, Theorem 1.3.21]. This certainly implies $N' = F_\ell - ir(F_\ell)$ to be diagonalizable (in particular, N' is a polynomial in F_ℓ). Thus, we have found a diagonalizable matrix $N' \in \mathcal{N}(B)$ with distance at most ε from N and the proof is complete. $\qquad\square$

REMARK 3.11. Notice that, if $N \in \mathcal{N}(B)$ in Theorem 3.10 is selfadjoint or skewadjoint, then in the proof of the theorem either $S = \frac{1}{2}(N + N^\star) = 0$ or $K = \frac{1}{2}(N - N^\star) = 0$ holds. Then the previous proof essentially reduces to the same reasoning used for the proof of Theorem 3.2. However, if N is B-unitary, the construction in proof of Theorem 3.10 does not necessarily give a matrix $N' \in \mathcal{N}(B)$ that is again B-unitary.

Applying Remark 2.1(c) and using the result of Theorem 3.10, we may now prove the density of diagonalizable matrices in $\mathcal{N}(B)$ without the assumption of B being unitary.

THEOREM 3.12. Let $B \in \mathbf{Gl}_n(\mathbb{C})$ with $B = B^H$. For any $N \in \mathcal{N}(B)$ and any $\varepsilon > 0$, there exists some diagonalizable $N' \in \mathcal{N}(B)$ such that $\|N - N'\|_2 \le \varepsilon$.

Proof. According to Theorem 3.10 the statement is true if B is unitary, so assume $B \in \mathbf{Gl}_n(\mathbb{C})$ is not unitary. According to [9, Theorem 4.5.7], there exists some $Q \in \mathbf{Gl}_n(\mathbb{C})$ such that

$$B' := Q^H B Q = \begin{bmatrix} -I_{m_-} & \\ & I_{m_+} \end{bmatrix},$$

where m_- (m_+) is the number of negative (positive) eigenvalues of B. According to Remark 2.1(c), we have $Q^{-1}\mathcal{N}(B)Q = \mathcal{N}(B')$. Since B' is unitary, Theorem 3.10 applies and the set of diagonalizable matrices is dense in $\mathcal{N}(B')$. As any matrix $A \in \mathcal{N}(B')$ is diagonalizable if and only if $QAQ^{-1} \in \mathcal{N}(B)$ is diagonalizable the density result follows for $\mathcal{N}(B)$. $\quad\square$

In Sections 3.1, 3.2 and 3.3, we showed the density results of semisimple matrices for $B = B^H$. Hereby, B was an arbitrary (indefinite) Hermitian matrix. Before we pass on to the case $B = -B^H$, do not overlook the special case of B being positive definite (and $[\cdot, \cdot]_B$ defining a scalar product). Due to Remark 2.1(c), the same situation as for $B = I_n$ takes place and all matrices in $\mathbb{J}(B), \mathbb{L}(B), \mathbb{G}(B)$ and $\mathcal{N}(B)$ are semisimple. In fact, for a Hermitian positive definite matrix $B \in \mathbf{Gl}_n(\mathbb{C})$ there exists some $Q \in \mathbf{Gl}_n(\mathbb{C})$ such that $Q^H BQ = I_n$. Whenever $A \in \mathbb{J}(B)$, then $A' := Q^{-1}AQ \in \mathbb{J}(I_n)$ is Hermitian and, in consequence, semisimple. As semisimplicity is preserved under similarity transformation, A must have been semisimple, too. The same reasoning holds in an analogous way for $\mathbb{L}(B), \mathbb{G}(B)$ and $\mathcal{N}(B)$ using Remark 2.1(c).

3.4. Skew-Hermitian sesquilinear forms. We now consider the case where $B = -B^H \in \mathbf{Gl}_n(\mathbb{C})$ is a skew-Hermitian matrix[4]. Fortunately, this situation can be completely traced back to the analysis from Sections 3.1, 3.2 and 3.3.

First note that, if $B = -B^H$ holds, then iB is Hermitian (i.e. $(iB)^H = -iB^H = iB$). Moreover, we have

$$(iB)^{-1}A^H(iB) = -iB^{-1}A^H(iB) = -i^2 B^{-1}A^H B = B^{-1}A^H B.$$

This shows that $A \in \mathbb{J}(B)$ ($A \in \mathbb{L}(B)$, resp.) if and only if $A \in \mathbb{J}(iB)$ ($A \in \mathbb{L}(iB)$, resp.). Moreover

$$A^H(iB)A = iB \iff i(A^H BA) = iB \iff A^H BA = B,$$

so $A \in \mathbb{G}(B)$ if and only if $A \in \mathbb{G}(iB)$. Finally, the same reasoning reveals that $\mathcal{N}(B) = \mathcal{N}(iB)$. Therefore, some matrix in any of these sets corresponding to B can always be interpreted as a matrix from the same set corresponding to iB. The main theorems obtained in the previous sections thus apply directly when $B = -B^H$.

THEOREM 3.13. *Let $B \in \mathbf{Gl}_n(\mathbb{C})$ with $B = -B^H$. For any $A \in \mathbf{M}_n(\mathbb{C})$ in any of the sets $\mathbb{J}(B), \mathbb{L}(B),$ $\mathbb{G}(B)$ and $\mathcal{N}(B)$ and any $\varepsilon > 0$ there exists some diagonalizable $A' \in \mathbf{M}_n(\mathbb{C})$ belonging to the same set such that $\|A - A'\|_2 \leq \varepsilon$. In addition, the set of matrices with pairwise distinct eigenvalues is dense in $\mathbb{J}(B),$ $\mathbb{L}(B)$ and $\mathbb{G}(B)$.*

Proof. Any matrix $A \in \mathbf{M}_n(\mathbb{C})$ from $\mathbb{J}(B), \mathbb{L}(B), \mathbb{G}(B)$ or $\mathcal{N}(B)$ can be interpreted as a matrix from $\mathbb{J}(iB), \mathbb{L}(iB), \mathbb{G}(iB)$ or $\mathcal{N}(iB)$, respectively, and Theorems 3.2, 3.7 and 3.12 apply. $\quad\square$

Certainly, there are analogous results for $B = -B^H$ as stated in Corollary 3.5 and Proposition 3.6.

4. Conclusions. In this work, we considered the structure classes of B-selfadjoint, B-skewadjoint, B-unitary and B-normal matrices defined by an (indefinite) scalar product $[x, y] = x^H By$ on $\mathbb{C}^n \times \mathbb{C}^n$ for some $B \in \mathbf{Gl}_n(\mathbb{C})$. We showed that, if $B = \pm B^H$, the set of semisimple (i.e. diagonalizable) matrices is dense in the set of all B-normal matrices. Additionally, we proved the density result for the sets of B-selfadjoint, B-skewadjoint and B-unitary matrices.

---

[4]Notice that n needs to be even for $B = -B^H$ to be nonsingular.

REFERENCES

[1] A.V. Arkhangel'skii and L.S. Pontryagin (editors). *General Topology I.* Springer Berlin, Heidelberg, 1990.

[2] R.J. De la Cruz and P. Saltenberger *Generic Canonical Forms for Perplectic and Symplectic Normal Matrices.* Preprint ArXiv 2006.16790v1, 2020.

[3] K. Fan and A.J. Hoffmann. Some metric inequalities in the space of matrices. *Proc. Am. Math. Soc.*, 6:111–116, 1955.

[4] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications.* Birkhäuser Verlag, Basel, 2005.

[5] G.H. Golub and C.F. Van Loan. *Matrix Computations* (3rd edition). Johns Hopkins University Press, Baltimore, 1996.

[6] G.J. Groenewald, D.B. Janse van Rensburg, and A.C.M. Ran. A canonical form for H-unitary matrices. *Oper. Matrices,* 10:739–783, 2016.

[7] R. Grone, C.R. Johnson, E.M. Sa, and H. Wolkowicz. Normal matrices. *Linear Algebra Appl.,* 87:213–225, 1987.

[8] D.J. Hartfiel. Dense sets of diagonalizable matrices. *Proc. Am. Math. Soc.,* 123:1669–1672, 1995.

[9] R.A. Horn and C.R. Johnson. *Matrix Analysis* (2nd edition). Cambridge University Press, New York, 2013.

[10] N. Jacobson. *Basic Algebra 1* (2nd Edition). H. W. Freeman and Company, New York, 1985.

[11] P. Lancaster and L. Rodman. *Algebraic Riccati Equations.* Clarendon Press, Oxford, 1995.

[12] D.W. Lewis. *Matrix Theory.* World Scientific Publishing, Singapore, 1991.

[13] D.S. Mackey, N. Mackey, and F. Tisseur. On the definition of two natural classes of scalar product. *MIMS Eprint 2007.64, Manchester Institute for Mathematical Sciences*, Manchester, 2007.

[14] C. Mehl. On classification of normal matrices in indefinite inner product spaces. *Electron. J. Linear Algebra,* 15(1):50–83, 2006.

[15] C. Mehl, V. Mehrmann, A.C.M. Ran, and L. Rodman. Eigenvalue perturbation theory of classes of structured matrices under generic structured rank one perturbations. *Linear Algebra Appl.,* 435:687–716, 2011.

[16] C. Mehl, V. Mehrmann, A.C.M. Ran, and L. Rodman. Perturbation theory of selfadjoint matrices and sign characteristics under generic structured rank one perturbations. *Linear Algebra Appl.,* 436:4027–4042, 2012.

[17] C. Mehl, V. Mehrmann, A.C.M. Ran, and L. Rodman. Eigenvalue perturbation theory of structured real matrices and their sign characteristics under generic structured rank-one perturbations. *Linear and Multilinear Algebra,* 64:527–556, 2016.

[18] K.C. O'Meara, J. Clark, and C.I. Vinsonhaler. *Advanced Topics in Linear Algebra.* Oxford University Press, Oxford, 2011.

[19] P. Saltenberger. *On Different Concepts for the Linearization of Matrix Polynomials and Canonical Canonical Decompositions of Structured Matrices with Respect to Indefinite Sesquilinear Forms.* Logos Verlag, Berlin, 2019.