

When is a Hamiltonian matrix the commutator of two skew-Hamiltonian matrices?

Vanni Noferini*

March 27, 2014

The mapping $\Phi_n(A, B) = AB - BA$, where the matrices $A, B \in \mathbb{C}^{2n \times 2n}$ are skew-Hamiltonian with respect to transposition, is studied. Let \mathcal{C}_n be the range of Φ_n : we give an implicit characterization of \mathcal{C}_n , obtaining results that find an application in algebraic geometry. Namely, they are used in [R. Abuaf and A. Boralevi, Orthogonal bundles and skew-Hamiltonian matrices, In Preparation] to study orthogonal vector bundles. We also give alternative and more explicit characterizations of \mathcal{C}_n for $n \leq 3$. Moreover, we prove that for $n \geq 4$ the complement of \mathcal{C}_n is nowhere dense in the set of $2n$ -dimensional Hamiltonian matrices, denoted by \mathcal{H}_n , implying that almost all matrices in \mathcal{H}_n are in \mathcal{C}_n for $n \geq 4$. Finally, we show that Φ_n is never surjective as a mapping from $\mathcal{W}_n \times \mathcal{W}_n$ to \mathcal{H}_n , where \mathcal{W}_n is the set of $2n$ -dimensional skew-Hamiltonian matrices. Along the way, we discuss the connections of this problem with several existing results in matrix theory.

Keywords: Hamiltonian matrix, skew-Hamiltonian matrix, commutator, orthogonal vector bundle, bow tie form, Roth's theorem

MSC classification: 15A21, 15B57

1 Introduction

It is a classical result in linear algebra that a square matrix with elements in a field can be written as the commutator of two matrices if and only if it is traceless [2]. This paper addresses a structured version of the same question for matrices with elements in the

*School of Mathematics, University of Manchester, Manchester, England, M13 9PL (vanni.noferini@manchester.ac.uk). This work was supported by ERC Advanced Grant MATFUN (267526). This preprint (printed with a different latex style) was submitted to Linear and Multilinear Algebra.

complex field. Clearly, if A and B are skew-Hamiltonian¹ then $AB - BA$ is Hamiltonian. An analysis of the reverse implication is less obvious: when can a complex Hamiltonian matrix H be written as the commutator of two skew-Hamiltonian matrices A and B ? In spite of the relative simplicity of the question, we find that the answer deserves some discussion.

Besides the interest that it may have, per se, as a structured linear algebra question, the problem is relevant to algebraic geometry as well. More precisely, it arises in the context of the study of the (possible) irreducibility of the moduli space of vector bundles on the projective plane with stable rank r and Chern classes $(0, n)$. The unstructured case of this problem is studied in [12], while its version for symplectic vector bundles is analyzed in [16]. The much more technically involved case of orthogonal vector bundles is addressed in [1]. In fact, the connection with algebraic geometry is the main motivation of our study, that originated after some discussions [3] with the author of [16] and one author of [1]. This paper can be considered both preliminary and complementary to [1, Sections 4-5]. Indeed, it is preliminary because some of our results are amongst the ingredients for the irreducibility theorem proved in [1, Theorem 3.4]. On the other hand, it is complementary as our point of view is different than [1] and we cover some aspects of the problem that are not considered there. Some results in our Section 4 and in [1, Section 5] are similar to each other, but we emphasize that they have been independently proved with different tools, providing different insights.

In [1], a certain mapping naturally arises. This mapping sends an ordered pair of complex skew-symmetric matrices of even dimension $2n$, say (X, Y) , to the complex symmetric matrix $S = XJY - YJX$, with $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. For a fixed value of n a characterization of the range of the mapping is of interest. As $A = JX$ and $B = JY$ are skew-Hamiltonian matrices of size $2n$, a natural observation is that we may equivalently consider the map Φ_n that sends (A, B) to the Hamiltonian matrix $AB - BA = JS$. After recalling some basic notions in Section 2, we follow this latter approach and in Section 3 we study the range of Φ_n , denoted by \mathcal{C}_n . Having formulated the problem in terms of Hamiltonian and skew-Hamiltonian matrices, the invariance under symplectical similarity of the problem becomes manifest, allowing us to exploit some results on structured canonical forms of skew-Hamiltonian matrices [6, 7]. Furthermore, our approach unveils several connections with similar problems in structured linear algebra, such as the existence of Hamiltonian square roots of skew-Hamiltonian matrices [13, 6, 7, 18] or other properties of skew-Hamiltonian, Hamiltonian and symplectic matrices [5, 13, 15, 18]. Section 4 analyzes how the topological properties of \mathcal{C}_n change from $n \leq 3$ to $n \geq 4$, as in the former case \mathcal{C}_n is a closed nowhere dense set, while in the latter case *the complement* of \mathcal{C}_n is contained in a closed nowhere dense set. This had led to the conjecture [3] that Φ_n could be surjective for $n \geq 4$. We exhibit a counterexample in Section 5. Finally, in Section 6 we discuss how this problem is related to Sylvester-like equations and Roth's theorem [17]. In particular, we observe that a result of [19] can be modified when dealing with Sylvester equations with skew-symmetry constraints.

¹In this paper, (skew)-Hamiltonianity is always defined with respect to transposition, not with respect to conjugate transposition. See Definition 2.1.

2 Basic facts on Hamiltonian, skew-Hamiltonian and symplectic matrices

Let $J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{C}^{2n \times 2n}$. Observe that J has the property $J^T = J^{-1} = -J$, that we will repeatedly use throughout the paper.

Definition 2.1. • A matrix $W \in \mathbb{C}^{2n \times 2n}$ is called *skew-Hamiltonian* if $W^T J = JW$. We denote the set of $2n$ -dimensional skew-Hamiltonian complex matrices by \mathcal{W}_n .

- A matrix $H \in \mathbb{C}^{2n \times 2n}$ is called *Hamiltonian* if it satisfies the equation $JH + H^T J = 0$. We denote the set of $2n$ -dimensional Hamiltonian complex matrices by \mathcal{H}_n .
- A matrix $V \in \mathbb{C}^{2n \times 2n}$ is called *symplectic* if $V^T J V = J$. We denote the set of $2n$ -dimensional symplectic complex matrices by \mathcal{S}_n .

Although some authors, e.g. [7], distinguish between two possible definitions of complex (skew-)Hamiltonian and symplectic matrices (that is, either with respect to transposition or with respect to conjugate transposition), in this paper we always consider, without ambiguity, the former choice. The three sets \mathcal{W}_n , \mathcal{H}_n and \mathcal{S}_n correspond to, respectively, the Jordan algebra, the Lie algebra and the automorphism group associated with the bilinear form $(x, y) \in \mathbb{C}^{2n} \times \mathbb{C}^{2n} \mapsto x^T J y \in \mathbb{C}$. Further discussions on this remark, together with a generalization to any nondegenerate bilinear or sesquilinear form over the complex field, can be found in [15].

Clearly, (skew-)Hamiltonian matrices and (skew-)symmetric matrices of even size are related. We omit the straightforward proof of the next lemma.

Lemma 2.2. $H \in \mathcal{H}_n$ if and only if $H = JS$, with $S = S^T \in \mathbb{C}^{2n \times 2n}$. $W \in \mathcal{W}_n$ if and only if $W = JK$, with $K = -K^T \in \mathbb{C}^{2n \times 2n}$.

Lemma 2.2 implies that Hamiltonian matrices have the form

$$\begin{bmatrix} M & F \\ G & -M^T \end{bmatrix} \quad (1)$$

with $F = F^T$ and $G = G^T$, while skew-Hamiltonian matrices can be written as

$$\begin{bmatrix} E & C \\ D & E^T \end{bmatrix} \quad (2)$$

with $C = -C^T$ and $D = -D^T$.

The next lemma recapitulates some spectral properties of matrices belonging to the three classes defined above.

Lemma 2.3. *The following facts hold:*

1. if $H \in \mathcal{H}_n$, λ is an eigenvalue of H if and only if $-\lambda$ is (as a corollary, if 0 is an eigenvalue then it must have even algebraic multiplicity);

2. in the Jordan canonical form of H , there is an even number of nilpotent Jordan blocks of odd size $2m + 1$, for all $m \in \mathbb{N}$;
3. if $W \in \mathcal{W}_n$, all its eigenvalues have even algebraic multiplicity;
4. if $V \in \mathcal{S}_n$, V is nonsingular and λ is an eigenvalue of V if and only if λ^{-1} is (as a corollary, $\det V = 1$).

Proof. 1. The eigenvalues of H are the roots of $\det(H - xI) = \det(H^T - xI) = \det(JHJ - xI) = \det(H + xI)$.

2. This can be seen, e.g., as a direct consequence of [14, Theorem 3.10] applied to the T -even matrix pencil $JH - Jx$.
3. Any skew-Hamiltonian matrix can be expressed as the square of a Hamiltonian matrix [7, Proposition 14]: see also [18, Theorem 5] for a generalization to any field of characteristic other than 2. Hence there exists a Hamiltonian matrix whose eigenvalues, if squared, give the eigenvalues of W .
4. The nonsingularity of V follows immediately from the definition. Moreover, $V = -JV^{-T}J$, so V and V^{-T} are similar, implying the second statement. \square

Given $V \in \mathcal{S}_n$, let us consider the operator $\Psi_V : \mathbb{C}^{2n \times 2n} \rightarrow \mathbb{C}^{2n \times 2n}$, $\Psi_V(X) := VXV^{-1}$. The structure-preserving action of the symplectic similarity operator Ψ_V is summarized in the following proposition.

Proposition 2.4. *Let $U, V \in \mathcal{S}_n$, $H \in \mathcal{H}_n$ and $W \in \mathcal{W}_n$. Then $\Psi_V(H) \in \mathcal{H}_n$, $\Psi_V(W) \in \mathcal{W}_n$ and $\Psi_V(U) \in \mathcal{S}_n$.*

Proof. The first two statements can be easily verified using $JH^T + HJ = 0$, $W^T J = JW$, and $-JV^T J = V^{-1}$. From the definition of \mathcal{S}_n one can check that symplectic matrices are a group under matrix multiplication. This fact implies the third statement. \square

Finally, we recall a remarkable result first proved in [6, Theorem 1] over the real field (and considering the real Jordan canonical form). As noted in [7], a simple modification of the original proof allows to state the following version over the complex field, which has appeared in [7, Theorem 4] and in [13, Theorem 3]:

Theorem 2.5. *Let $B \in \mathcal{W}_n$. There exists $V \in \mathcal{S}_n$ such that*

$$\Psi_V(B) = \begin{bmatrix} J_B & 0 \\ 0 & J_B^T \end{bmatrix}$$

and J_B is in Jordan canonical form, hence unique up to permutation of Jordan blocks.

We denote by \mathcal{J}_n the set of $2n$ -dimensional skew-Hamiltonian matrices that are in the canonical form of Theorem 2.5.

3 On commutators of skew-Hamiltonian matrices

The commutator of two matrices $X, Y \in \mathbb{C}^{m \times m}$ is defined as $[X, Y] := XY - YX$. It is immediate to check that if $X, Y \in \mathcal{W}_n$ then $[X, Y] \in \mathcal{H}_n$. Consider the map

$$\Phi_n : \mathcal{W}_n \times \mathcal{W}_n \rightarrow \mathcal{H}_n, \quad \Phi_n(A, B) := [A, B].$$

The map Φ_n naturally arises in algebraic geometry, when orthogonal vector bundles are investigated [1]. In particular, in that context it is of interest to describe the range of Φ_n , that we denote by \mathcal{C}_n . In the present section we give a characterization of \mathcal{C}_n that finds an important application in [1], where it is used in order to prove the irreducibility of the moduli space of stable rank r orthogonal vector bundles on \mathbb{P}^2 , with Chern classes $(0, n)$ and trivial splitting on the general line.

For small values of n , one can convince oneself that Φ_n is not surjective. Here are some claims that we are going to prove in Section 4. When $n = 1$, $\mathcal{C}_n = \{0\}$. For $n = 2$, the eigenvalues of a matrix in \mathcal{C}_n must be $\{\lambda, \lambda, -\lambda, -\lambda\}$ for some $\lambda \in \mathbb{C}$; for $n = 3$, the spectrum of a matrix in \mathcal{C}_n must be of the form $\{\lambda, \mu, \lambda + \mu, -\lambda, -\mu, -\lambda - \mu\}$ for some $\lambda, \mu \in \mathbb{C}$. As the dimension grows to $n \geq 4$, no constraint on the spectrum is any longer imposed, with the obvious exception of the $\lambda \leftrightarrow -\lambda$ symmetry that, as Lemma 2.3 illustrates, is induced by the Hamiltonian structure. Furthermore, \mathcal{C}_n can be shown to contain an open dense subset of \mathcal{H}_n for $n \geq 4$. From these observations, a natural conjecture is that [3] Φ_n could be surjective when n is large enough. This leads to the following problem:

Problem 3.1. *For $n \geq 4$, is $\mathcal{C}_n = \mathcal{H}_n$?*

The answer turns out to be negative. In Section 5, we will provide a counterexample, thus proving that the inclusion $\mathcal{C}_n \subset \mathcal{H}_n$ is strict for all n .

A first step towards a characterization of \mathcal{C}_n comes from observing that Proposition 2.4 shows that Ψ_V acts as an automorphism on both \mathcal{W}_n and \mathcal{H}_n . Moreover, we have:

Proposition 3.2. $\Psi_V(\Phi(A, B)) \equiv \Phi(\Psi_V(A), \Psi_V(B))$.

Proof. By direct verification. □

Let us define \mathcal{R}_n as the range of the restriction of Φ_n to $\mathcal{W}_n \times \mathcal{J}_n$. In view of Theorem 2.5, Proposition 2.4 and Proposition 3.2, it is clear that

$$\mathcal{C}_n = \{X \in \mathcal{H}_n \mid \exists V \in \mathcal{S}_n \text{ s.t. } \Psi_V(X) \in \mathcal{R}_n\}.$$

With this in mind, we now investigate the structure of \mathcal{R}_n .

Proposition 3.3. *If $B \in \mathcal{J}_n$ is as in Theorem 2.5, and writing $A \in \mathcal{W}_n$ as in (2) then denoting by $\text{Sym}(X) = (X + X^T)/2$:*

$$\Phi_n(A, B) = \begin{bmatrix} [E, J_B] & \text{Sym}(-2J_B C) \\ \text{Sym}(2DJ_B) & -([E, J_B])^T \end{bmatrix}.$$

Proof. It follows from a direct computation, observing that $C = -C^T$, that $D = -D^T$, and that $([E, J_B])^T = -[E^T, J_B^T]$. \square

The problem is thus reduced to three coupled subproblems:

1. When can a matrix M be written as $[E, J_B]$ where E is any matrix while J_B is in Jordan form?
2. When can a symmetric matrix F be written as the symmetric part of $-2J_B C$ where C is skew-symmetric and J_B is the same as in the subproblem above?
3. When can a symmetric matrix G be written as the symmetric part of $2DJ_B$ where D is skew-symmetric and J_B is the same as the two subproblems above?

In the following subsections we will characterize the properties of any triple of matrices M , F and G solving the three subproblems, thus characterizing \mathcal{R}_n . To this goal, a few definitions come handy.

Definition 3.4. Let $A \in \mathbb{C}^{m \times m}$. If $|k| < m$ the k th diagonal of A is the set $\{A_{ij} \mid j - i = k\}$. The k th trace of A is defined as the sum of the elements in the k th diagonal of A . The k th antidiagonal of A is the set $\{A_{ij} \mid i + j = m + 1 - k\}$. The sum of the elements in the k th antidiagonal of A is called the k th antitrace of A .

Definition 3.5. Let $A \in \mathbb{C}^{m \times m}$. We say that A is subtraceless if its k th traces are zero for all $k \leq 0$. We say that A is supertraceless if its k th traces are zero for all $k \geq 0$. We say that A is subantitraceless if its k th antitraces are zero for all $k \leq 0$. We say that A is superantitraceless if its k th antitraces are zero for all $k \geq 0$.

Lemma 3.6. Let N be the $m \times m$ nilpotent Jordan block and $R \in \mathbb{C}^{m \times m}$ with $R_{ij} = \delta_{i+j, m+1}$. Then, for any $A \in \mathbb{C}^{m \times m}$ and any $0 \leq k \leq m - 1$, the following holds: (i) the $(-k)$ th trace of A is $\text{tr}(N^k A)$; (ii) the k th trace of A is $\text{tr}(N^k R A R)$; (iii) the $(-k)$ th antitrace of A is $\text{tr}(R N^k A)$; (iv) the k th antitrace of A is $\text{tr}(N^k R A)$.

Proof. We have $\text{tr}(N^k A) = \sum_{i,j=1}^m (N^k)_{ji} A_{ij} = \sum_{j=1}^{m-k} \sum_{i=k+1}^m \delta_{i,j+k} A_{ij} = \sum_{i=k+1}^m A_{i,i-k}$, that by definition is the $(-k)$ th trace. This proves statement (i). We omit the analogous proof of statements (ii)-(iv). \square

3.1 Subproblem 1

Suppose that J_B has ℓ Jordan blocks. Denote by Υ_i its i th Jordan block, and suppose Υ_i has eigenvalue λ_i and size d_i , with $\sum_{i=1}^{\ell} d_i = n$. Partitioning both E and M coherently with the size of the Jordan blocks, we get the Sylvester equations

$$M_{ij} = E_{ij} \Upsilon_j - \Upsilon_i E_{ij}.$$

For $i \neq j$ each equation describes a Sylvester linear endomorphism of $\mathbb{C}^{d_i \times d_j}$ mapping E_{ij} to M_{ij} . It is known [10, Theorem 2.4.4.1] that such an endomorphism is bijective if and

only if $\lambda_i \neq \lambda_j$. In order to analyze the diagonal blocks of M , we need an appropriate variation on a theme of [6]. A special case of [6, Proposition 4] states that if Υ is a $d \times d$ Jordan block then $A \mapsto \Upsilon A - A\Upsilon^T$ is a surjective map from $\mathbb{C}^{d \times d}$ to the space of subantitraceless matrices of size d .

Lemma 3.7. *Let Υ be any $d \times d$ Jordan block and $\phi_1(A) = [\Upsilon, A]$. Then ϕ_1 is a surjective mapping from $\mathbb{C}^{d \times d}$ to the space of subtraceless square matrices of size d .*

Proof. Write $\Upsilon = \lambda I_d + N$ where N is a nilpotent Jordan block. As $[\Upsilon, A] = [N, A]$, there is no loss of generality in taking $\lambda = 0$. For any $k \geq 0$, by the linearity of the trace, $\text{tr}(N^k \phi_1(A)) = \text{tr}(N^{k+1}A) - \text{tr}(N^k AN) = 0$. Thus, by Lemma 3.6, the range of ϕ_1 is included in the space of subtraceless matrices of size d , which is the kernel of the linear map $X \in \mathbb{C}^{d \times d} \mapsto [\text{tr}(N^0 X) \quad \dots \quad \text{tr}(N^{d-1} X)]^T \in \mathbb{C}^d$. This map is clearly surjective, as $[v \quad 0 \quad \dots \quad 0] \mapsto v \in \mathbb{C}^d$, and hence, subtraceless matrices are a $d^2 - d$ dimensional subspace of $\mathbb{C}^{d \times d}$. It is a classical result [8, Chapter VIII] that $\ker \phi_1$ is the space of upper triangular Toeplitz matrices of size d , which has dimension d over \mathbb{C} . Thus, the rank-nullity theorem implies that ϕ_1 is surjective. \square

3.2 Subproblem 2

We look for a skew-symmetric solution C to the equation $J_B C - C J_B^T = F$. Partition C and F coherently with the dimension of the Jordan blocks in J_B . We get

$$F_{ij} = \Upsilon_i C_{ij} - C_{ij} \Upsilon_j^T.$$

The off-diagonal blocks are once again associated with a map that is bijective if and only if $\lambda_i \neq \lambda_j$. To analyze the diagonal blocks, we need another auxiliary result. Although the mapping that we study in the following lemma is a special case of the one analyzed in [6, Proposition 4], an important detail is that here we restrict its domain to the set of skew-symmetric matrices of size d .

Lemma 3.8. *Let Υ be any $d \times d$ Jordan block. Then the map $\phi_2(Z) = Z\Upsilon^T - \Upsilon Z$ is a bijection from the space of d -dimensional skew-symmetric matrices to the space of d -dimensional symmetric subantitraceless matrices. In particular if $d = 1$ then $\phi_2(0) = 0$, while if $d \geq 2$ the three elements in the right-bottom corner of $\phi_2(Z)$ must be zero.*

Proof. We can assume without loss of generality $\Upsilon = N$ where N is the nilpotent Jordan block of size d . The symmetry of $\phi_2(Z)$ is manifest from the definition. Moreover $XN^T - NX$ is subantitraceless for any $X \in \mathbb{C}^{d \times d}$ [6, Proposition 4]. Now we show that ϕ_2 is a bijection. Skew-symmetric matrices are a $\frac{d(d-1)}{2}$ -dimensional vector space over \mathbb{C} , and so are subantitraceless symmetric matrices. To show this, since symmetric matrices have dimension $\frac{d(d+1)}{2}$, we invoke Lemma 3.6 and argue that the map $X = X^T \mapsto [\text{tr}(RN^0 X) \quad \dots \quad \text{tr}(RN^{d-1} X)]^T \in \mathbb{C}^d$ is surjective. Indeed, take any $\alpha \in \mathbb{C}$ and

$w \in \mathbb{C}^{d-1}$, and observe $\begin{bmatrix} 0 & w \\ w^T & \alpha \end{bmatrix} \mapsto \begin{bmatrix} 2w \\ \alpha \end{bmatrix}$. Hence, it suffices to prove that ϕ_2 is injective. We induct on d . When $d = 1$, ϕ_2 maps the trivial vector space $\{0\}$ to itself. Now partition

$$Z = \begin{bmatrix} 0 & v^T \\ -v & Z_* \end{bmatrix}, \quad N = \begin{bmatrix} 0 & e_1^T \\ 0 & N_* \end{bmatrix} \quad \Rightarrow \quad \phi_2(Z) = \begin{bmatrix} 2v^T e_1 & v^T N_*^T - e_1^T Z_* \\ Z_* e_1 + N_* v & Z_* N_*^T - N_* Z_* \end{bmatrix}.$$

Here $Z_* = -Z_*^T \in \mathbb{C}^{(d-1) \times (d-1)}$, N_* is the nilpotent Jordan block of size $d-1$, $v \in \mathbb{C}^{d-1}$, and $e_1 \in \mathbb{C}^{d-1}$ with $(e_1)_i = \delta_{i,1}$. Suppose $\phi_2(Z) = 0$. Looking at the right-bottom block in $\phi_2(Z)$, we get $Z_* = 0$ by the inductive hypothesis. Substituting in the remaining block elements of $\phi_2(Z)$ we see that $v = 0$ as well. \square

3.3 Subproblem 3

We finally aim to find a skew-symmetric solution D to the equation $DJ_B - J_B^T D = G$. Partition D and G coherently with the dimension of the Jordan blocks in J_B . We get the equations

$$G_{ij} = D_{ij} \Upsilon_j - \Upsilon_i^T D_{ij}$$

Everything is akin to the previous section, modulo a reflection with respect to the 0th antidiagonal.

Lemma 3.9. *Let Υ be any $d \times d$ Jordan block. Then the map $\phi_3(Z) = Z\Upsilon - \Upsilon^T Z$ is a bijection from the space of d -dimensional skew-symmetric matrices to the space of d -dimensional symmetric superantitraceless matrices. In particular if $d = 1$ then $\phi_3(0) = 0$, while if $d \geq 2$ the three elements in the top-left corner of X must be zero.*

Proof. Analogous to the proof of Lemma 3.8. \square

3.4 Coupling the three subproblems

We now bring all the pieces together. A few definitions will help to expose the results.

Definition 3.10. *A partition of a positive integer $n > 1$ is an ordered ℓ -uple of positive integers d_1, \dots, d_ℓ such that $d_1 + \dots + d_\ell = n$. We denote the class of the partitions of n as π_n .*

Definition 3.11. *If $\beta = (d_1, \dots, d_\ell) \in \pi_n$, then the β -block partition of a matrix $X \in \mathbb{C}^{n \times n}$ is the set of blocks X_{ij} , $i, j = 1, \dots, \ell$, such that $X_{ij} \in \mathbb{C}^{d_i \times d_j}$ and*

$$X = \begin{bmatrix} X_{11} & \dots & X_{1\ell} \\ \vdots & & \vdots \\ X_{\ell 1} & \dots & X_{\ell\ell} \end{bmatrix}.$$

Definition 3.12. *A Hamiltonian matrix $H = \begin{bmatrix} M & F \\ G & -M^T \end{bmatrix} \in \mathcal{H}_n$ is said to be in bow tie form if there exists $\beta = (d_1, \dots, d_\ell) \in \pi_n$ such that the diagonal blocks in the β -block partitions of M , F and G satisfy the following:*

- M_{ii} are subtraceless for all $i = 1, \dots, \ell$;
- F_{ii} are subantitraceless for all $i = 1, \dots, \ell$;
- G_{ii} are superantitraceless for all $i = 1, \dots, \ell$.

Figure 1: Matrices in bow tie form with $n = 4$ and $\beta = (4)$ (left) and $\beta = (1, 1, 1, 1)$ (right).

\star $\star \ \star$ $\star \ \star \ \star$ $0 \ \star \ \star \ \star$	 $\bullet \ \bullet \ 0$ $\bullet \ \bullet \ 0 \ 0$	0 0 0 0	0 0 0 0
$0 \ 0 \ \bullet \ \bullet$ $0 \ \bullet \ \bullet$ $\bullet \ \bullet$ \bullet	$\star \ \star \ \star \ 0$ $\star \ \star \ \star$ $\star \ \star$ \star	0 0 0 0	0 0 0 0

To justify the name, here is a visual representation of a matrix in bow tie form with $n = 4$ and $\beta = (4)$ (Figure 1, left). The symbol \star denotes elements that belong to a diagonal whose corresponding trace is zero; the symbol \bullet denotes elements that belong to an antidiagonal whose corresponding antitrace is zero; elements that need to be zero are explicitly indicated, whereas all other elements are not drawn. The Hamiltonian structure is also tacitly assumed, but not depicted. With the same interpretation of symbols (but admittedly much less resembling a bow tie), Figure 1 also displays, on the right, the other extreme of a matrix in bow tie form with $n = 4$ and $\beta = (1, 1, 1, 1)$.

We mention that in [1], where the problem of interest is stated in terms of symmetric and skew-symmetric matrices, a symmetric matrix S such that JS is in bow tie form is said to be in *diamond form*.

We can now prove the following theorem.

Theorem 3.13. $H \in \mathcal{R}_n$ if and only if H is in bow tie form.

Proof. Let $H \in \mathcal{R}_n$. Then by definition there exist $A \in \mathcal{W}_n$ and $B = \begin{bmatrix} J_B & 0 \\ 0 & J_B^T \end{bmatrix} \in \mathcal{J}_n$ such that $H = [A, B]$. Let J_B have Jordan blocks of sizes d_1, \dots, d_ℓ . By Lemmata 3.7, 3.8, 3.9, necessarily H is in bow tie form, and the associated partition of n is $\beta = (d_1, \dots, d_\ell)$. Note that, if J_B is derogatory, then further constraints appear for some off-diagonal blocks, corresponding to those Sylvester linear maps that fail to be bijective. Yet, H must still be in bow tie form, as Lemmata 3.7, 3.8, 3.9 do not depend on the eigenvalues of J_B .

Conversely, suppose that H is in bow tie form for some $\beta = (d_1, \dots, d_\ell) \in \pi_n$. Our goal is to find $A \in \mathcal{W}_n$ partitioned as in (2) and $B \in \mathcal{J}_n$ partitioned as in Theorem 2.5 such that $H = [A, B]$. Let us take any nonderogatory J_B in Jordan form and with ℓ Jordan

blocks of sizes d_1, \dots, d_ℓ . Then Lemmata 3.7, 3.8, 3.9, that characterize three surjective maps, together with classical results on the nonsingularity of the Sylvester operator in the disjoint spectrum case, guarantee that A can be found such that $H = [A, B]$. \square

Given $H \in \mathcal{R}_n$, associated with $\beta \in \pi_n$, and as long as J_B is a nonderogatory matrix in Jordan form with blocks of sizes prescribed by β , Theorem 3.13 guarantees that there is always $A = \begin{bmatrix} E & C \\ D & E^T \end{bmatrix} \in \mathcal{W}_n$ such that $[A, J_B \oplus J_B^T] = H$. Here and elsewhere in the paper \oplus denotes a direct sum. We observe that the blocks C and D in A are uniquely determined by H and by J_B ; on the other hand, the off-diagonal blocks in the β -partition of E are uniquely determined while the diagonal blocks are only defined up to any matrix in $\ker \phi_1$, i.e., any upper triangular Toeplitz matrix.

We are now ready to give a first answer to our initial question.

Corollary 3.14. *Let $H \in \mathcal{H}_n$. Then $H \in \mathcal{C}_n$ if and only if H is symplectically similar to a Hamiltonian matrix in bow tie form.*

Proof. The result follows from Theorem 3.13, Theorem 2.5, and Proposition 3.2. \square

3.5 Further remarks

Corollary 3.14 implicitly describes \mathcal{C}_n , and is used in [1] to tackle a nontrivial problem in algebraic geometry. Yet, it is possible to strengthen the statement of Corollary 3.14. To this goal, we first need to recall some definitions.

Definition 3.15. [9, Chapter 1] *A square root of $A \in \mathbb{C}^{n \times n}$ is a matrix X such that $X^2 = A$. If X is a square root of A , it is called a primary square root of A if $X = f(A)$ for some polynomial f . In this case we write $X = \sqrt{A}$.*

A primary square root of A exists if and only if $\text{rank } A = \text{rank } A^2$: see, e.g., [11, Theorem 6.4.12 (b)]. In particular, any nonsingular matrix has a primary square root. Singular skew-Hamiltonian matrices may not have skew-Hamiltonian square roots [7, Proposition 14]. Yet, every nonsingular skew-Hamiltonian matrix has a skew-Hamiltonian primary square root.

Lemma 3.16. *Any nonsingular $W \in \mathcal{W}_n$ has a primary square root $U \in \mathcal{W}_n$.*

Proof. Let $V \in \mathcal{S}_n$ be such that $V^{-1}WV = J_B \oplus J_B^T$ is in the form 2.5, and note that by assumption J_B has no zero eigenvalues. Let $\sqrt{J_B}$ be any primary square root of J_B . Then $U = V^{-1}(\sqrt{J_B} \oplus \sqrt{J_B^T})V$ is a primary square root of W by [9, Theorem 1.13 (b-c)]. Moreover, it is skew-Hamiltonian by Proposition 2.4. \square

Following [5], for $M \in \mathbb{C}^{2n \times 2n}$ we now define $M^* = -JM^TJ$. It is easy to check that $(MN)^* = N^*M^*$ and $(M^*)^* = M$. Note moreover that for $H \in \mathcal{H}_n$ it holds $H^* = -H$, while for $W \in \mathcal{W}_n$ one has $W^* = W$, and finally for $S \in \mathcal{S}_n$ one gets $S^*S = I_{2n}$. If M is invertible, we denote $M^{-*} = (M^*)^{-1} = (M^{-1})^*$. Finally, observe that, if f is any matrix

function (in the linear algebraic sense [9]), then $f(M^*) = f(-JM^T J) = -Jf(M^T)J = -Jf(M)^T J = f(M)^*$.

The following theorem was proved in [5, Theorem 4].

Theorem 3.17. *Let $H_0, H_1 \in \mathcal{H}_n$ and $W_0, W_1 \in \mathcal{W}_n$. Suppose that the pencils $L_0(x) = H_0 + xW_0$ and $L_1(x) = H_1 + xW_1$ are strictly equivalent. Then there exists a nonsingular matrix M such that $H_1 = MH_0M^*$ and $W_1 = MW_0M^*$.*

Furthermore, if P and Q are nonsingular and such that $L_1(x) = PL_0(x)Q$, then in the equations above one can take $M = P\sqrt{P^{-1}Q^}$, i.e., M can be taken to be equal to P times a primary square root of the nonsingular matrix $P^{-1}Q^*$.*

Proof. For completeness, we summarize the proof given in [5]. By $PH_0Q = H_1$ we obtain $H_1^* = Q^*H_0^*P^*$, hence $H_1 = Q^*H_0P^*$. Let $U = QP^{-*}$. Then from the equations above we easily get $U^*H_0 = H_0U$. Hence, for any polynomial f , $f(U^*)H_0 = H_0f(U)$. If $f(U)$ is nonsingular, this is equivalent to $f(U^*)H_0f(U)^{-1} = H_0$, implying $H_1 = Pf(U^*)H_0f(U)^{-1}Q$.

Similarly, we have $PW_0Q = W_1$, thus $W_1 = Q^*W_0P^*$. Arguing as before we get $W_1 = Pf(U^*)W_0f(U)^{-1}Q$. Let $M = Pf(U^*)$, and let us consider the functional equation $M^* = f(U)^{-1}Q$. We readily obtain $f(U)^2 = U$. As U is nonsingular, one can always choose $f(U)$ to be a primary square root of U . Then $M = P\sqrt{P^{-1}Q^*}$. \square

Theorem 3.17 has the following consequence.

Theorem 3.18. *Any two Hamiltonian matrices are similar if and only if they are symplectically similar.*

Proof. One implication is trivial. Assume now $H_1 = S^{-1}H_0S$ for some invertible matrix S . Let $M = S^{-1}\sqrt{SS^*}$ where the square root is primary (for example, as $SS^* \in \mathcal{W}_n$ is nonsingular, one can follow the explicit construction of Lemma 3.16). Applying Theorem 3.17 to $(H_1 - xI_{2n}) = S^{-1}(H_0 - xI_{2n})S$ we obtain $MM^* = I_{2n}$, that is, $M \in \mathcal{S}_n$. Furthermore, $H_1 = MH_0M^* = MH_0M^{-1}$. \square

Theorem 3.18 has a counterpart for skew-Hamiltonian matrices [13, Theorem 5]. We emphasize that Theorem 3.18 considers similarity *over the complex field*. Its statement is not true over the real field, unlike its analogue for skew-Hamiltonian matrices, that holds over any field of characteristic not 2 [18, Theorem 4]. For example, the two real

Hamiltonian matrices $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ are similar but *not* symplectically similar over

\mathbb{R} . However, by Theorem 3.18, they are symplectically similar over \mathbb{C} , e.g., via $\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$.

The same example also illustrates that Theorem 3.18 no longer holds when we consider Hamiltonianity and symplecticity with respect to conjugate transposition, as opposed to transposition as in Definition 2.1.

The main result of this section is that, for a Hamiltonian matrix, having the same Jordan form of some matrix in bow tie form is equivalent to being the commutator of two skew-Hamiltonian matrices.

Corollary 3.19. *Let $H \in \mathcal{H}_n$. Then $H \in \mathcal{C}_n$ if and only if H is similar to a Hamiltonian matrix in bow tie form.*

Proof. The statement follows from Corollary 3.14 and Theorem 3.18. \square

4 On the set \mathcal{C}_n

In this section we investigate further the different topological properties of the set \mathcal{C}_n for small and large values of n . Although in this paper we do not pursue this view, the problem, given $H \in \mathcal{H}_n$, to find $A, B \in \mathcal{W}_n$ such that $H = AB - BA$ can be regarded as a system of $2n^2 + n$ independent polynomial equations in $4n^2 - 2n$ unknowns. The question of interest is for which choice of the $2n^2 + n$ constant terms in the equations the system is consistent, i.e., admits at least one solution. Clearly, the system is underdetermined for $n \geq 2$. At a first glance, one might expect the system to be consistent for $n \geq 2$ and a generic choice of H . A rigorous analysis reveals a subtler situation. Indeed, a phase transition occurs: \mathcal{C}_n is nowhere dense in \mathcal{H}_n when $n \leq 3$, whereas its complement is nowhere dense for $n \geq 4$. Here we follow a matrix theoretic approach; [1, Proposition 5.1] provides an analysis from the point of view of algebraic geometry, recovering similar results by means of the different tools.

4.1 $n \leq 3$ revisited

As any pair of 2×2 skew-Hamiltonian matrix commute, it is obvious that \mathcal{C}_1 has only one element: the zero matrix. When $n = 1$, Corollary 3.19 is just another way to state the previous observation: the only matrix in bow tie form is $H = 0$.

When $n = 2$, either $\beta = (2)$ or $\beta = (1, 1)$. The corresponding parametric families of matrices in bow tie form are, respectively, $\begin{bmatrix} a & b & c & 0 \\ 0 & -a & 0 & 0 \\ 0 & 0 & -a & 0 \\ 0 & d & -b & a \end{bmatrix}$ and $\begin{bmatrix} 0 & a & 0 & c \\ b & 0 & c & 0 \\ 0 & d & 0 & -b \\ d & 0 & -a & 0 \end{bmatrix}$.

By squaring the two matrices above we obtain, respectively, $a^2 I_4$ and $(ab + cd)I_4$. Since $[\Psi_V(H)]^2 = \Psi_V(H^2)$, we easily see that $\Psi_V(H^2) = \kappa^2 I_4 \Leftrightarrow H^2 = \kappa^2 I_4$ for some $\kappa \in \mathbb{C}$. In other words, \mathcal{C}_2 is included in the set of Hamiltonian square roots of the identity times a scalar. In fact, the two sets coincide.

Proposition 4.1. $\mathcal{C}_2 = \{H \in \mathcal{H}_2 \mid \exists \kappa \in \mathbb{C} \text{ s.t. } H^2 = \kappa^2 I_4\}$

Proof. Let $H \in \mathcal{C}_2$. Then, as argued above, by direct inspection $H^2 = \kappa^2 I_4$ for some κ . We now prove that this condition is also sufficient. Invoking Corollary 3.19, it suffices to show that any Hamiltonian square root of $H^2 = \kappa^2 I_4$ is similar to a matrix in bow tie form. Arguing on the Jordan decomposition of H we see that if $\kappa \neq 0$ then H must be diagonalizable with exactly two eigenvalues equal to κ and exactly two eigenvalues equal to $-\kappa$. Therefore, it is (symplectically) similar to the first form above with $a = \kappa$, $b = c = d = 0$. If $\kappa = 0$, then the only eigenvalue is 0 and 2 is an upper bound on the length of a Jordan chain. There are three possibilities. The case of $H = 0$ is trivial. If the Jordan form of H has one Jordan block of dimension 2 and two Jordan blocks of dimension 1, H is similar to the first form above with $c = 1$, $a = b = d = 0$. Finally, if

the Jordan form of H has two Jordan blocks of dimension 2, H is similar to the first form above with $b = 1$, $a = c = d = 0$. \square

To analyze \mathcal{C}_3 , the following technical lemma is useful.

Lemma 4.2. *Let $a, c \in \mathbb{C}$ and denote by x_1, x_2, x_3 the three solutions of the equation $x^3 + 2cx^2 + c^2x + a = 0$. Then there is a choice of the signs of their square roots such that $\sqrt{x_1} = \sqrt{x_2} + \sqrt{x_3}$.*

Proof. Let σ denote the sign function, extended to complex numbers as follows: $\sigma(0) = 0$; if $z \in i\mathbb{R}$ then $\sigma(z) = \sigma(\Im(z))$; in all other instances, $\sigma(z) = \sigma(\Re(z))$. Define $s(z)$ to be the unique square root of z s.t. $\sigma(s(z)) \neq -1$. Note that, for any z , $\sigma(z)s(z) = s(z)^2$. Now, define $\mu := \sigma(x_1 - x_2 - x_3)$ and $\nu := \sigma(s(x_2) + \mu s(x_3))$. The equation $(x_1 + x_2 + x_3)^2 = 4(x_1x_2 + x_2x_3 + x_3x_1)$ can be rewritten as $(x_1 - x_2 - x_3)^2 = 4x_2x_3$. This yields $x_1 - x_2 - x_3 = 2\mu s(x_2)s(x_3)$. Thus, $s(x_1) = s(x_2 + x_3 + 2\mu s(x_2)s(x_3)) = \nu s(x_2) + \mu\nu s(x_3)$. Therefore, it suffices to pick $\sqrt{x_1} = s(x_1)$, $\sqrt{x_2} = \nu s(x_2)$ and $\sqrt{x_3} = \mu\nu s(x_3)$. \square

There are four possible partitions of 3: $(1, 1, 1)$, $(2, 1)$, (3) and $(1, 2)$. Here are the parametric families of matrices in bow tie form corresponding to the former three:

$$\begin{bmatrix} 0 & a & b & 0 & c & d \\ e & 0 & f & c & 0 & j \\ g & h & 0 & d & j & 0 \\ 0 & k & \ell & 0 & -e & -g \\ k & 0 & m & -a & 0 & -h \\ \ell & m & 0 & -b & -f & 0 \end{bmatrix}; \begin{bmatrix} a & b & c & d & 0 & e \\ 0 & -a & f & 0 & 0 & g \\ h & j & 0 & e & g & 0 \\ 0 & 0 & k & -a & 0 & -h \\ 0 & \ell & m & -b & a & -j \\ k & m & 0 & -c & -f & 0 \end{bmatrix}; \begin{bmatrix} a & b & c & d & e & f \\ g & h & j & e & -2f & 0 \\ 0 & -g & -a-h & f & 0 & 0 \\ 0 & 0 & k & -a & -g & 0 \\ 0 & -2k & m & -b & -h & g \\ k & m & \ell & -c & -j & a+h \end{bmatrix}.$$

A fourth parametric family, not shown here, corresponds to the partition $(1, 2)$: neglecting it is not a loss of generality, because it is permutation similar to the second parametric family above. The following result characterizes \mathcal{C}_3 .

Proposition 4.3. *$H \in \mathcal{C}_3$ if and only if its eigenvalues are of the form $\{\lambda, \mu, \lambda + \mu, -\lambda, -\mu, -\lambda - \mu\}$ for some $\lambda, \mu \in \mathbb{C}$.*

Proof. Let H be in any of the three forms above. Because of the Hamiltonian structure, it must hold $\det(xI - H) = x^6 + \alpha x^4 + \beta x^2 + \gamma$, where $\alpha, \beta, \gamma \in \mathbb{Z}[a, b, c, d, e, f, g, h, j, k, \ell, m]$. By direct computation it can be checked that, for each of the three forms above, $\alpha^2 = 4\beta$ must hold: we omit the lengthy, but straightforward, calculations. By Lemma 4.2, this is equivalent to the necessity of the spectral condition in the statement. Note that any other matrix in \mathcal{R}_3 is permutation similar to one of the three forms above, hence has the same characteristic polynomial.

To prove the reverse implication, we show that every Jordan form compatible with this spectral condition and with the Hamiltonian structure is similar to a matrix realized by a clever choice of the parameters in the third parametric family displayed above. Due to the restriction on the number of odd-size nilpotent Jordan blocks (see Proposition 2.3), there are nine possible cases. In the table below, λ and μ are any pair of complex numbers, and $J_k(\lambda)$ is the Jordan block of size k and eigenvalue λ . For each case, we prescribe a corresponding choice of parameters, with the convention that all parameters not explicitly mentioned in the corresponding table entry are implicitly set to be 0. It is easy to check

that each corresponding Jordan form is achieved by these choices: we omit the details. We emphasize that λ, μ are allowed to be 0. Hence, some of the entries of the table simultaneously describe cases where H has different rank. For example, the Jordan form $J_2(0) \oplus J_2(0) \oplus 0 \oplus 0$ can be achieved setting $\lambda = 0$ in either the second or the third entry

	Jordan form of H	Values of parameters
	$\lambda \oplus \mu \oplus (\lambda + \mu) \oplus -\lambda \oplus -\mu \oplus (-\lambda - \mu)$	$a = \lambda, h = \mu$
	$J_2(\lambda) \oplus J_2(-\lambda) \oplus 2\lambda \oplus -2\lambda$	$a = h = \lambda, b = 1$
	$J_2(\lambda) \oplus J_2(-\lambda) \oplus 0 \oplus 0$	$a = -h = \lambda, e = 1$
	$J_2(\lambda) \oplus J_2(-\lambda) \oplus J_2(0)$	$a = -h = \lambda, e = \ell = 1$
of the table.	$\lambda \oplus \lambda \oplus -\lambda \oplus -\lambda \oplus J_2(0)$	$a = -h = \lambda, \ell = 1$
	$J_3(0) \oplus J_3(0)$	$b = j = 1$
	$J_4(0) \oplus 0 \oplus 0$	$k = 1, b = d = 2$
	$J_4(0) \oplus J_2(0)$	$d = k = 1$
	$J_6(0)$	$g = d = 1$

4.2 $n \geq 4$ revisited

We have just seen how, for $n \leq 3$, matrices in \mathcal{C}_n must satisfy certain additional constraints on their spectrum. The situation changes as the dimension reaches $n \geq 4$, as the following proposition shows.

Proposition 4.4. *For $n \geq 4$, given any matrix in \mathcal{H}_n there exists a matrix in \mathcal{C}_n with the same spectrum. In particular, the set of Hamiltonian matrices with all distinct eigenvalues is included in \mathcal{C}_n .*

Proof. The idea is to show that we can always find a matrix in \mathcal{R}_n having any prescribed spectrum compatible with the Hamiltonian structure. This implies, in particular, that if $H \in \mathcal{H}_n$ has all distinct eigenvalues then $H \in \mathcal{C}_n$. Indeed, in this case there is only one possible Jordan form and there is nothing else to check. We start with $n = 4$. Let

$$H = \begin{bmatrix} 0 & F \\ G & 0 \end{bmatrix} \text{ with}$$

$$F = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 & 1 & \chi \\ 0 & -2 & -\chi & 0 \\ 1 & -\chi & \gamma_2/2 & -\gamma_1/4 \\ \chi & 0 & -\gamma_1/4 & \gamma_0/2 \end{bmatrix}.$$

Evidently, $H \in \mathcal{R}_4$, associated with $\beta = (4)$, so $H \in \mathcal{C}_4$. Moreover, it holds $\det(xI_8 - H) = \det(x^2I_4 - FG)$. By a direct computation $\det(FG + (y - \chi)I_4) = y^4 + \gamma_2y^2 + \gamma_1y + \gamma_0$. Suppose that we wish to fix the eigenvalues of FG , whose square roots give the eigenvalues of H , to be a certain unordered quadruple of complex numbers $\{x_1, x_2, x_3, x_4\}$. Since the map $\{x_1, x_2, x_3, x_4\} \mapsto (\chi, \gamma_0, \gamma_1, \gamma_2)$ is bijective, we can always choose G accordingly. Similar constructions exist for $n = 5, 6, 7$. For example, for $n = 5$ let $\omega_{\pm} := \frac{-1 \pm i\sqrt{15}}{4}$ and

set:

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & \omega_+ \\ 0 & 0 & 0 & \omega_- & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \omega_- & 0 & 0 & 0 \\ \omega_+ & 0 & 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 & 0 & 1 & \chi\omega_- \\ 0 & 0 & -1 & \chi\omega_+ & 0 \\ 0 & -1 & \chi & \gamma_3\omega_+/2 & 0 \\ 1 & \chi\omega_+ & \gamma_3\omega_+/2 & \gamma_2\omega_+^2 & -\gamma_1\omega_+/2 \\ \chi\omega_- & 0 & 0 & -\gamma_1\omega_+/2 & \gamma_0 \end{bmatrix}.$$

Again, F is subantitraceless, G is superantitraceless, and it can be checked that $\det(FG + (y - \chi)I_5) = y^5 + \sum_{j=0}^3 \gamma_j y^j$. For $n = 6$ an analogous argument is generated by considering $\omega_{\pm} := \frac{-1 \pm i\sqrt{3}}{2}$ and:

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & \omega_+ \\ 0 & 0 & 0 & \omega_- & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \omega_- & 0 & 0 & 0 \\ \omega_+ & 0 & 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 & 0 & 0 & 1/2 & \chi\omega_- \\ 0 & 0 & 0 & -1 & \chi\omega_+ & 0 \\ 0 & 0 & 1 & \chi & 0 & 0 \\ 0 & -1 & \chi & -\gamma_4 & -\gamma_3\omega_+/2 & 0 \\ 1/2 & \chi\omega_+ & 0 & -\gamma_3\omega_+/2 & -\gamma_2\omega_- & \gamma_1\omega_+ \\ \chi\omega_- & 0 & 0 & 0 & \gamma_1\omega_+ & -4\gamma_0 \end{bmatrix}.$$

For $n = 7$, we can again define $\omega_{\pm} := \frac{-1 \pm i\sqrt{15}}{4}$ and set:

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \omega_+ \\ 0 & 0 & 0 & 0 & \omega_- & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \omega_- & 0 & 0 & 0 & 0 \\ \omega_+ & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1/2 & \chi\omega_- \\ 0 & 0 & 0 & 0 & -1 & \chi\omega_+ & 0 \\ 0 & 0 & 0 & 1/2 & \chi & 0 & 0 \\ 0 & 0 & 1/2 & -\chi & \gamma_5 & 0 & 0 \\ 0 & -1 & \chi & \gamma_5 & -4\gamma_4 & -2\gamma_3\omega_+ & 0 \\ 1/2 & \chi\omega_+ & 0 & 0 & -2\gamma_3\omega_+ & -4\gamma_2\omega_+^2 & 4\gamma_1\omega_+ \\ \chi\omega_- & 0 & 0 & 0 & 0 & 4\gamma_1\omega_+ & -16\gamma_0 \end{bmatrix}.$$

Finally, when $n \geq 8$, we can take F and G to be block diagonal, with the size of each block being between 4 and 7. \square

We now equip \mathcal{H}_n with the metric topology inherited by its embedding in the normed space $(\mathbb{C}^{2n \times 2n}, \|\cdot\|)$. Here $\|\cdot\|$ is any suitable matrix norm, e.g., the Frobenius norm $\|A\|_F = \sqrt{\sum_{i,j} |A_{ij}|^2}$. Proposition 4.4 implies that almost all matrices in \mathcal{H}_n are in \mathcal{C}_n when $n \geq 4$.

Proposition 4.5. *If $n \geq 4$, then \mathcal{C}_n contains an open dense subset of \mathcal{H}_n .*

Proof. Let $N = 2n^2 + n$. We define a linear map $\mathcal{M} : \mathcal{H}_n \rightarrow \mathbb{C}^N$ by rearranging (in an arbitrary, but fixed, order) an arbitrary, but fixed, set of N independent entries of $H \in \mathcal{H}_n$ into a vector $\mathcal{M}(H) \in \mathbb{C}^N$. The coefficients of $p = \det(H - xI_{2n})$ are polynomials in the independent entries of H . By Proposition 4.4, $H \notin \mathcal{C}_n \Rightarrow \mathcal{M}(H) \in \Omega$, having defined $\Omega \subset \mathbb{C}^N$ to be the algebraic variety $\text{Res}(p, p') = 0$, where Res is the resultant. Since Ω is Zariski closed, it is a closed nowhere dense subset of \mathbb{C}^N in the Euclidean topology. As \mathcal{M} is a homeomorphism, $\mathcal{M}^{-1}(\mathbb{C}^N \setminus \Omega)$ is open and dense in \mathcal{H}_n , implying the statement because $\mathcal{M}^{-1}(\mathbb{C}^N \setminus \Omega) \subseteq \mathcal{C}_n$. \square

An alternative way to state the latter result is that $\mathcal{H}_n \setminus \mathcal{C}_n$ is contained in a closed nowhere dense set, and hence is nowhere dense. It is interesting to compare our proof with the alternative approach, employing different techniques, of [1, Proposition 5.1].

Although Proposition 4.5 may sound like a clue in favour of an affirmative answer to Problem 3.1, it turns out that the complement of \mathcal{C}_n is not the empty set, no matter how large n is. Proving this fact is the scope of the next section.

5 Solution of Problem 3.1

Our argument is based on the construction of a concrete counterexample. We start from two auxiliary lemmata and a technical result to be used in the following.

Lemma 5.1. *Let $S \in \mathbb{C}^{n \times n}$ be such that $\text{rank } S \leq 2$ and $S = S^T$. Then $S = aa^T + bb^T$ for some vectors a and b .*

Proof. It is a corollary of the Autonne-Takagi factorization of a symmetric matrix over the complex field [10, Corollary 2.6.6 (a)]. \square

Lemma 5.2. *Let $H = Jaa^T + Jbb^T \in \mathcal{H}_n$, with $a, b \in \mathbb{C}^{2n}$. Then the characteristic polynomial of H is $x^{2n} + \Theta_n^2 x^{2n-2}$, where $\Theta_n = a^T Jb = \sum_{j=1}^n (a_j b_{n+j} - b_j a_{n+j})$.*

Proof. As H is Hamiltonian and has at most two nonzero eigenvalues, it suffices to prove the equation for Θ_n . We have $2\Theta_n^2 = \text{tr}(H)^2 - \text{tr}(H^2) = -\text{tr}(H^2)$, observing that Hamiltonian matrices are traceless. From the cyclic property of the trace, $\text{tr}(H^2) = \text{tr}(M^2)$, where $M = \begin{bmatrix} a^T \\ b^T \end{bmatrix} J \begin{bmatrix} a & b \end{bmatrix}$. Since $v^T Jv = 0 \forall v \in \mathbb{C}^{2n}$, we deduce that $M = \begin{bmatrix} 0 & a^T Jb \\ -a^T Jb & 0 \end{bmatrix}$. Hence, $\Theta_n^2 = -\text{tr}(M^2)/2 = (a^T Jb)^2$. \square

Proposition 5.3. *Let $u, v \in \mathbb{C}^s$ and $X = uu^T + vv^T$. If X is superantitraceless, then, for some $\eta \in \{i, -i\}$ and some $r \in \{0, 1, \dots, \lceil \frac{s}{2} \rceil\}$, the following conditions hold:*

1. $u_j = v_j = 0 \forall j \leq r$;
2. if $r < \lceil \frac{s}{2} \rceil$ then $u_{r+1} \neq 0$ and $v_j = \eta u_j \forall r+1 \leq j \leq s-r$.

If X is subantitraceless, then, for some $\eta \in \{i, -i\}$ and some $q \in \{0, 1, \dots, \lceil \frac{s}{2} \rceil\}$, the following conditions hold:

1. $u_j = v_j = 0 \forall j \geq s+1-q$;
2. if $q < \lceil \frac{s}{2} \rceil$ then $u_{s-q} \neq 0$ and $v_j = \eta u_j \forall q+1 \leq j \leq s-q$.

Proof. Let $R = R^T \in \mathbb{C}^{s \times s}$ have elements $R_{ij} = \delta_{i+j, s+1}$. For any $v \in \mathbb{C}^s$, Rv has the same elements of v , but in the reverse order. Similarly, if $X \in \mathbb{C}^{s \times s}$ then RXR has the same elements of X , but after having flipped both the rows and the columns. Therefore, $X = uu^T + vv^T$ is subantitraceless if and only if $RXR = Ruu^T R^T + Rvv^T R^T =$

$(Ru)(Ru)^T + (Rv)(Rv)^T$ is superantitraceless. Hence, applying the first part of the statement to RXR we obtain the second part.

It remains to prove the first part. Using Lemma 3.6, the $(s-j)$ th antitrace of the matrix $X = uu^T + vv^T$ can be written as $\tau_j = \text{tr}(N^{s-j}RX) = u^T N^{s-j}Ru + v^T N^{s-j}Rv$,

where $N \in \mathbb{C}^{s \times s}$ is a nilpotent Jordan block. Hence, $\tau_j = \sum_{k=1}^j u_k u_{j+1-k} + v_k v_{j+1-k}$. The superantitraceless condition is equivalent to $\tau_j = 0$ for $j = 1, \dots, s$. If $u = 0$, define $r' = s$; otherwise, let $0 \leq r' < s$ be such that $u_j = 0 \forall j \leq r'$ but $u_{r'+1} \neq 0$.

Suppose first that $0 \leq 2r' < s$. By definition, this implies $u_{r'+1} \neq 0$. We argue that $v_j = 0 \forall j \leq r'$. If $r' = 0$, this is a vacuous claim and there is nothing to prove. Otherwise $0 = \tau_1 = v_1^2$, yielding $v_1 = 0$. Now suppose that $v_j = 0 \forall j \leq \ell < r'$. With this assumption, $\tau_{2\ell+1} = v_{\ell+1}^2$. As $2\ell + 1 < 2r' + 1 \leq s$, we deduce that $v_{\ell+1} = 0$. This completes an argument by strong induction that shows that $v_1 = \dots = v_{r'} = 0$. The latter condition guarantees that $\tau_j = 0$ for all $j \leq 2r'$. Furthermore, it yields $\tau_{2r'+1} = u_{r'+1}^2 + v_{r'+1}^2 = 0$, implying $v_{r'+1} = \eta u_{r'+1}$ with $\eta^2 = -1$. Suppose now $v_j = \eta u_j$ (with the same choice of the sign of η for all j) for all $r' + 1 \leq j \leq \ell < s - r'$.

Then $\tau_{r'+\ell+1} = \sum_{k=r'+1}^{\ell+1} u_k u_{r'+\ell+2-k} + v_k v_{r'+\ell+2-k} = u_{r'+1}(u_{\ell+1} + \eta v_{\ell+1}) = 0$. Hence, $v_{\ell+1} = \eta u_{\ell+1}$, and again by strong induction we conclude that the statement of the Lemma is true with $r = r'$.

Finally let $2r' \geq s$ and let $\rho = \lceil \frac{s}{2} \rceil$. An argument akin to the one above shows that $v_1 = \dots = v_\rho = 0$: note that, in this case, the latter condition implies $\tau_j = 0$ for all $j \leq s \leq 2\rho$, i.e., it is also sufficient for X to be superantitraceless. Choosing $r = \rho$, the proof is completed. Indeed, if $r = \lceil \frac{s}{2} \rceil$ then the premise of condition (2) is false so there is nothing else to check. \square

We are now ready for the main result of this section.

Theorem 5.4. *Let $\theta \neq 0$. If $H \in \mathcal{H}_n$ is diagonalizable and has eigenvalues $\{\theta, -\theta, \underbrace{0, \dots, 0}_{2n-2 \text{ times}}\}$,*

then $H \notin \mathcal{C}_n$.

Proof. Suppose $H \in \mathcal{C}_n$. Then there exists $L \in \mathcal{R}_n$ similar to H . Thus, $\text{rank } L = 2$, yielding $L = J(aa^T + bb^T)$ by Proposition 2.2 and Lemma 5.1. Hence, $\det(xI - L) = x^{2n} + \Theta_n^2 x^{2n-2}$ by Lemma 5.2, and thus $\Theta_n^2 = -\theta^2$. We argue that $L \in \mathcal{R}_n$ implies $\Theta_n = 0$, contradicting the assumptions.

Let $\beta = (d_1, \dots, d_\ell)$ be the partition of n associated with the particular bow tie form of L . Let $\Theta_{n,k} := \sum_{j=d_1+\dots+d_{k-1}+1}^{d_1+\dots+d_{k-1}+d_k} (a_j b_{n+j} - b_j a_{n+j})$. We claim that $\Theta_{n,k} = 0$ for all $k = 1, \dots, \ell$, and thus $\Theta_n = \sum_{k=1}^{\ell} \Theta_{n,k} = 0$.

We now prove that $\Theta_{n,1} = \sum_{j=1}^{d_1} (a_j b_{n+j} - b_j a_{n+j}) = 0$. Generalizing the argument to show $\Theta_{n,k} = 0 \forall k$ is straightforward. Denote $s := d_1$ and $a_u := [a_1 \ \dots \ a_s]^T$, $b_u :=$

$[b_1 \ \dots \ b_s]^T, a_d := [a_{n+1} \ \dots \ a_{n+s}]^T, b_d := [b_{n+1} \ \dots \ b_{n+s}]^T \in \mathbb{C}^s$. With the notation of Definition 3.12, let us consider the blocks M_{11}, F_{11}, G_{11} in L . Then we have $G_{11} = -a_u a_u^T - b_u b_u^T$, $F_{11} = a_d a_d^T + b_d b_d^T$, and $M_{11} = a_u a_d^T + b_u b_d^T$.

In particular, we can apply Proposition 5.3 to infer statements on the vectors a_u, b_u, a_d, b_d . First, since G_{11} (hence $-G_{11}$) must be superantitraceless, we conclude from Proposition 5.3 that there is some $r \leq \lceil \frac{s}{2} \rceil$ such that $a_j = b_j = 0 \ \forall 1 \leq j \leq r$. Furthermore, unless $r = \lceil \frac{s}{2} \rceil$, we also know that $a_{r+1} \neq 0$ and there is a choice of the sign of $\eta_1 = \pm i$ such that $b_j = \eta_1 a_j \ \forall r+1 \leq j \leq s-r$. Similarly, as F_{11} is subantitraceless, we deduce from Proposition 5.3 that there is some $q \leq \lceil \frac{s}{2} \rceil$ such that $a_{n+j} = b_{n+j} = 0$ for all $s-q+1 \leq j \leq s$. Moreover, if $q < \lceil \frac{s}{2} \rceil$, then it holds $a_{n+s-q} \neq 0$ and $b_{n+j} = \eta_2 a_{n+j} \ \forall q+1 \leq j \leq s-q$ for a choice of the sign in $\eta_2 = \pm i$.

By definition of r and q , we can write

$$\Theta_{n,1} = \sum_{j=r+1}^{s-q} (a_j b_{n+j} - b_j a_{n+j}). \quad (3)$$

Thus, we may assume $r+q < s$, as otherwise $\Theta_{n,1} = 0$ and there is nothing else to prove. We still have to consider the subtraceless property of $M_{11} = a_u a_d^T + b_u b_d^T$. Lemma 3.6 yields $0 = a_d^T N^{s-k} a_u + b_d^T N^{s-k} b_u$ for $1 \leq k \leq s$. Hence, taking into account the previous discussion, the nontrivial equations to be satisfied are

$$\sum_{j=r+1}^{k-q} (a_j a_{n+s+j-k} + b_j b_{n+s+j-k}) = 0 \quad \text{for } r+q+1 \leq k \leq s. \quad (4)$$

It is convenient to split the analysis into three cases, according to the sign of $r-q$.

$r = q$. We have $q = r < \frac{s}{2} \leq \lceil \frac{s}{2} \rceil$, implying both $a_{r+1} \neq 0$ and $a_{n+s-r} \neq 0$. Moreover (3) is simplified to

$$\Theta_{n,1} = \sum_{j=r+1}^{s-r} [a_j a_{n+j} (\eta_2 - \eta_1)]. \quad (5)$$

Note that for all $2r+1 \leq k \leq s$ and $r+1 \leq j \leq k-r$ it holds $s-r \geq s+j-k \geq j \geq r+1 \geq r+1+k-s$, so that (4) become $\sum_{j=r+1}^{k-r} [a_j a_{n+s+j-k} (1 + \eta_1 \eta_2)] = 0$. In particular, when $k = 2r+1$, we get $a_{r+1} a_{n+s-r} (1 + \eta_1 \eta_2) = 0$. Thus, $\eta_1 = \eta_2$. We conclude from (5) that $\Theta_{n,1} = 0$.

$r > q$. Since $q < \frac{s}{2} \leq \lceil \frac{s}{2} \rceil$, we have $a_{n+s-q} \neq 0$. There are two subcases.

If $2r < s$, then $r < \lceil \frac{s}{2} \rceil$, hence $a_{r+1} \neq 0$. Furthermore, from (3) we get

$$\Theta_{n,1} = \sum_{j=r+1}^{s-r} [a_j a_{n+j} (\eta_2 - \eta_1)] + \sum_{j=s-r+1}^{s-q} [a_{n+j} (\eta_2 a_j - b_j)]. \quad (6)$$

Arguing similarly to the $r = q$ case, we simplify (4) to

$$\sum_{j=r+1}^{k-q} [a_j a_{n+s+j-k} (1 + \eta_1 \eta_2)] = 0 \quad \text{for } r + q + 1 \leq k \leq s + q - r;$$

$$\sum_{j=r+1}^{s-r} [a_j a_{n+s+j-k} (1 + \eta_1 \eta_2)] + \sum_{j=s-r+1}^{k-q} a_{n+s+j-k} (a_j + \eta_2 b_j) = 0 \quad \text{for } s + q - r + 1 \leq k \leq s.$$

Setting $k = r + q + 1$ yields $a_{r+1} a_{n+s-q} (1 + \eta_1 \eta_2) = 0$, hence $\eta_1 = \eta_2$. Thus, all the equations in the first group are satisfied, while the equations in the second group become $\sum_{j=s-r+1}^{k-q} a_{n+s+j-k} (a_j + \eta_2 b_j) = 0$. When $k = s + q - r + 1$ we get $a_{n+s-q} (a_{s-r+1} + \eta_2 b_{s-r+1}) = 0$, implying $b_{s-r+1} = \eta_2 a_{s-r+1}$. Substituting into the next equations, we obtain by strong induction that $b_j = \eta_2 a_j$ for all $j \leq s - q$. It follows from (6) that $\Theta_{n,1}$ vanishes.

If $2r \geq s$, (3) becomes $\Theta_{n,1} = \sum_{j=r+1}^{s-q} a_{n+j} (\eta_2 a_j - b_j)$, whereas (4) become

$$\sum_{j=r+1}^{k-q} a_{n+s+j-k} (a_j + \eta_2 b_j) = 0 \quad \text{for } r + q + 1 \leq k \leq s.$$

When $k = r + q + 1$ we get $a_{n+s-q} (a_{r+1} + \eta_2 b_{r+1}) = 0$, yielding $b_{r+1} = \eta_2 a_{r+1}$. The now familiar strong induction argument lets us deduce that $b_j = \eta_2 a_j$ for all $j \leq s - q$, implying by (6) that $\Theta_{n,1} = 0$.

$r < q$. Mutatis mutandis, this case is completely analogous to the previous one: the reader is invited to fill in the details. \square

It is worth mentioning that Theorem 5.4 provides a counterexample only for finite values of n . Indeed, given the infinite diagonal matrix $D = 1 \oplus 0 \oplus 0 \dots$, consider the Hamiltonian operator $H = \begin{bmatrix} D & 0 \\ 0 & -D \end{bmatrix}$ and the skew-Hamiltonian operator $A = \begin{bmatrix} N & 0 \\ 0 & N^T \end{bmatrix}$ where N is an infinite nilpotent Jordan block. Then $H = [A, A^T]$.

5.1 Generalizations of the results and open questions

The problem discussed in this paper can be generalized in the following way. Let $S \in GL(m, \mathbb{C})$ and consider the nondegenerate bilinear form $(x, y) \mapsto x^T S y$. Define the set of S -Hamiltonian matrices as $\{H \in \mathbb{C}^{m \times m} \mid H^T S + S H = 0\}$ and the set of S -skew-Hamiltonian matrices as $\{A \in \mathbb{C}^{m \times m} \mid A^T S = S A\}$. When can an S -Hamiltonian matrix H be written as the commutator of two S -skew-Hamiltonian matrices A, B ? An analysis for a generic S will be presented in future work. Here we sketch how the problem can be tackled when $S = \pm S^T$. The results of this paper can be extended to any $S = -S^T \in GL(m, \mathbb{C})$ (i.e., the associated nondegenerate bilinear form is *alternating* [18]). Indeed, in this case $m = 2n$ must be even, and there exists $V \in GL(m, \mathbb{C})$ such that $S = V^T J V$: see, e.g., [18, p. 387] or [10, Corollary 2.6.6(b)]. Therefore, H is

S -Hamiltonian if and only if VHV^{-1} is J -Hamiltonian, and A is S -skew-Hamiltonian if and only if VAV^{-1} is J -skew-Hamiltonian. Hence, it is immediate to generalize the results obtained for $S = J$ to any alternating bilinear form. Much simpler is the case $S = S^T \in GL(m, \mathbb{C})$. In this instance, it follows from [10, Corollary 2.6.6(a)] that there exists $U \in GL(m, \mathbb{C})$ such that $S = U^T U$. Thus, it suffices to consider $S = I$, and it is easy to show that any skew-symmetric matrix can be written as the commutator of two symmetric matrices.

We now briefly discuss an open problem. Let $\mathcal{H}_{n,r}$ be the set of Hamiltonian matrices of size $2n$ and rank r . In the context of [1], another meaningful question is whether $\mathcal{H}_{n,r} \subseteq \mathcal{C}_n$ for a given value of r and n large enough. Theorem 5.4 shows that $\mathcal{H}_{n,2} \not\subseteq \mathcal{C}_n \forall n$. Using the results exposed in this paper, it is not difficult to verify that $\mathcal{H}_{n,1} \subseteq \mathcal{C}_n$ if and only if $n \geq 2$, while trivially $\mathcal{H}_{n,0} = \{0\} \subseteq \mathcal{C}_n \forall n$. The question remains open for $r \geq 3$.

6 Roth's theorem for Sylvester equations with skew-symmetry constraints

As argued in the introduction, the problem of writing a Hamiltonian matrix as the commutator of two skew-Hamiltonian matrices is closely related to the equation $YJX - XJY = C$ where $X = -X^T$, $Y = -Y^T$ and $C = C^T$. For fixed Y , this can be seen as a special case of the Sylvester matrix equation $AX - XB = C$, with the additional constraint $X = -X^T$. In this section, we briefly discuss this problem, relaxing the further constraints $B = A^T \in \mathcal{W}_n$ and $C = C^T$. In [19, Theorem 3], the celebrated Roth's theorem [17] is refined to give necessary and sufficient conditions for a Sylvester equation to have a Hermitian solution; moreover, the proof can be modified to deal with a symmetric solution as well. In our case, the solution is required to be skew-symmetric, and it turns out that an appropriate variation of [19, Theorem 3] can be stated.

Theorem 6.1. *Let \mathbb{F} be a field of characteristic different than 2. Let $A, B, C \in \mathbb{F}^{n \times n}$ and $F = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \in \mathbb{F}^{2n \times 2n}$. Then the Sylvester equation $AX - XB = C$ has a skew-symmetric solution $X = -X^T \in \mathbb{F}^{n \times n}$ if and only if there exists $S \in GL(2n, \mathbb{F})$ such that (a) $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} S = S \begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ and (b) $S^T F S = F$.*

Proof. The proof follows, with minor modifications, the arguments in [19]. If $X = -X^T$ is such that $AX - XB = C$, then it suffices to define $S = \begin{bmatrix} I_n & X \\ 0 & I_n \end{bmatrix}$. Conversely, suppose that an invertible S exists such that (a)-(b) hold. For any $Y \in \mathbb{F}^{n \times n}$ define $M_Y = \begin{bmatrix} A & Y \\ 0 & B \end{bmatrix}$ and consider the sets $\Omega_Y = \{Z \in \mathbb{F}^{2n \times 2n} \mid M_0 Z = Z M_Y \text{ and } F Z^T F M_0 = M_Y F Z^T F\}$. Manifestly, Ω_Y is a vector space over \mathbb{F} for any Y . Observe that $M_0(ZS) = (ZS)M_C \Leftrightarrow M_0 Z S = Z M_0 S \Leftrightarrow M_0 Z = Z M_0$, and that $F(ZS)^T F M_0 = M_C F(ZS)^T F \Leftrightarrow$

$S^{-1}FZ^TFM_0 = M_C S^{-1}FZ^TF \Leftrightarrow FZ^TFM_0 = M_0FZ^TF$. Hence, $Z \in \Omega_0 \Leftrightarrow ZS \in \Omega_C$, implying that $\dim \Omega_0 = \dim \Omega_C$.

Consider now the linear maps $\mathcal{M}_Y : \Omega_Y \rightarrow \mathbb{F}^{2n \times n}$, $Z = \begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix} \mapsto \mathcal{M}_Y(Z) = \begin{bmatrix} Z_1 \\ Z_3 \end{bmatrix}$. Comparing the block equations defining Ω_Y for $Y = 0$ and $Y = C$, it is easily checked that if $\begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix} \in \Omega_C$ then $\begin{bmatrix} Z_1 & 0 \\ Z_3 & 0 \end{bmatrix} \in \Omega_0$, implying that $\text{Im } \mathcal{M}_C \subseteq \text{Im } \mathcal{M}_0$. Moreover, again by definition of Ω_Y , $\ker \mathcal{M}_Y$ depends on A and B , but not on Y . Therefore, $\ker \mathcal{M}_C = \ker \mathcal{M}_0$. From the rank-nullity theorem $\dim \text{Im } \mathcal{M}_C = \dim \Omega_C - \dim \ker \mathcal{M}_C = \dim \Omega_0 - \dim \ker \mathcal{M}_0 = \dim \text{Im } \mathcal{M}_0$. Hence, $\text{Im } \mathcal{M}_C = \text{Im } \mathcal{M}_0$. Since clearly $I_{2n} \in \Omega_0$, this implies that Ω_C contains a matrix of the form $\begin{bmatrix} I_n & Z_2 \\ 0 & Z_4 \end{bmatrix}$. Substituting in the equations that define Ω_C , we get $AZ_2 - Z_2B = C$ and $Z_2^T B - AZ_2^T = C$. We conclude that $AX - XB = C$ for $X = (Z_2 - Z_2^T)/2$. \square

We mention that an analogue of Roth's theorem [4, Theorem 2.3] exists for the equation $AX + X^T B = C$, that is also related to the commutator problem.

7 Acknowledgements

I thank Giorgio Ottaviani and Ada Boralevi for sharing this problem with me and for several useful discussions, and Antonio Di Scala for his comments on the manuscript. The observation that Theorem 5.4 is not valid for infinite dimensional operators is due to Luca Gemignani.

References

- [1] R. Abuaf and A. Boralevi, Orthogonal bundles and skew-Hamiltonian matrices, in preparation. Preliminary version available at <http://arxiv.org/abs/1312.0963>
- [2] A. A. Albert and B. Muckenhoupt, On matrices of trace zero. Michigan Math. J. 4(1), 1–3, 1957.
- [3] A. Boralevi and G. Ottaviani, private communications with the author.
- [4] F. De Terán and F. Dopico, Consistency and efficient solution of the Sylvester equation for \star -congruence: $AX + X^*B = C$. Electron. J. Linear Algebra 22, 849–863, 2011.
- [5] H. Fassbender and K. D. Ikramov, Several observations on symplectic, Hamiltonian, and skew-Hamiltonian matrices. Linear Algebra Appl. 400, 15–29, 2005.
- [6] H. Fassbender, D. S. Mackey, N. Mackey and H. Xu, Hamiltonian square roots of skew-Hamiltonian Matrices. Linear Algebra Appl. 287(1–3), 125–129, 1999.

- [7] H. Fassbender, D. S. Mackey, N. Mackey and H. Xu, Hamiltonian square roots of skew-Hamiltonian Matrices. Extended version. <http://www.math.ku.edu/~xu/arch/FMMX.pdf>
- [8] F. R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea, Providence, RI, USA, 1960. (Translation from Russian).
- [9] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, PA, USA, 2008.
- [10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, NY, USA, second edition, 2013.
- [11] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, NY, USA, 1991.
- [12] K. Hulek. On the classification of stable rank- r vector bundles over the projective plane. In *Vector bundles and differential equations. Proceedings, Nice, France, June 12-17 1979*, editor A. Hirschowitz. Progr. Math., vol. 7, Birkäuser, Boston, MA, USA, 113–144, 1980.
- [13] K. D. Ikramov. Hamiltonian square roots of skew-Hamiltonian matrices revisited. *Linear Algebra Appl.* 325, 101–107, 2001.
- [14] D. S. Mackey, N. Mackey, C. Mehl and V. Mehrmann. Jordan structures of alternating matrix polynomials. *Linear Algebra Appl.* 432(4), 867–891, 2010.
- [15] D. S. Mackey, N. Mackey and F. Tisseur. Structured Factorizations in Scalar Product Spaces. *SIAM J. Matrix Anal. Appl.*, 27(3):821–850, 2006.
- [16] G. Ottaviani. Symplectic bundles on the plane, secant varieties and Lüroth quartics revisited. In *Vector bundles and low codimensional subvarieties: state of the art and recent developments*, editors G. Casnati, F. Catanese and R. Notari. Quad. Mat., vol. 21, Dept. Math. Seconda Univ. Napoli, Caserta, Italy, 315–352, 2007.
- [17] W. E. Roth, The Equations $AX - YB = C$ and $AX - XB = C$ in matrices. *Proc. Amer. Math. Soc.* 3, 392–396, 1952.
- [18] W. C. Waterhouse, The structure of alternating-Hamiltonian matrices. *Linear Algebra Appl.* 396, 385–390, 2005.
- [19] H. K. Wimmer. Roth’s Theorems for Matrix Equations With Symmetry Constraints. *Linear Algebra Appl.* 199, 357–362, 1994.