Products of positive definite symplectic matrices

by

Daryl Q. Granario

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 8, 2020

Keywords: symplectic matrix, positive definite matrix, matrix decomposition

Approved by

Tin-Yau Tam, Committee Chair and Professor Emeritus of Mathematics
Ming Liao, Professor of Mathematics
Huajun Huang, Associate Professor of Mathematics
Ulrich Albrecht, Department Chair and Professor of Mathematics

Abstract

We show that every symplectic matrix is a product of five positive definite symplectic matrices and five is the best in the sense that there are symplectic matrices which are not product of less.

In Chapter 1, we provide a historical background and motivation behind the study. We highlight the important works in the subject that lead to the formulation of the problem. In Chapter 2, we present the necessary mathematical prerequisites and construct a symplectic *congruence canonical form for $\mathrm{Sp}(2n, \mathbb{C})$. In Chapter 3, we give the proof of the main theorem. In Chapter 4, we discuss future research. The main results in this dissertation can be found in [18].

Acknowledgments

I would like to express my sincere gratitude to everyone who have supported me throughout this process. To my advisor Professor Tin-Yau Tam for the continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Thank you for their insightful comments and constant encouragement, and for the challenging questions that pushed me to grow as a mathematician. I thank Professor Ulrich Albrecht and Professor Ming Liao, who have been great mentors to me during my stay in Auburn. The lectures that they gave allowed me to have a deeper understanding and appreciation of mathematics beyond my specialization. I also thank Dr. Huajun Huang and Dr. Guofu Niu, who has provided good suggestions in improving the quality of this work.

Table of Contents

List of Tables

Chapter 1

Introduction

A lot of work has been done in expressing matrices and operators as products whose factors have "nice" properties. These properties of course, depends on the problem being tackled. The set of all normal matrices is an example of a collection of matrices frequently used and desired in applications. Let $M_n(\mathbb{C})$ be the set of $n \times n$ complex matrices. A matrix $A \in M_n(\mathbb{C})$ is *normal* if $AA^* = A^*A$, where $*$ is the conjugate transpose operator. Normal matrices are well-studied; a list of ninety equivalent conditions can be found in the surveys [11, 19]. Important subsets of normal matrices include the *unitary matrices* ($A^*A = I_n$), *Hermitian matrices* ($A^* = A$), and *skew-Hermitian matrices* ($A^* = -A$). A Hermitian matrix $A \in M_n(\mathbb{C})$ is *positive semidefinite* if $\sigma(A) \subseteq [0, \infty)$ and *positive definite* if $\sigma(A) \subseteq (0, \infty)$, where $\sigma(X)$ denotes the spectrum of $X \in M_n(\mathbb{C})$.

Every matrix $A \in M_n(\mathbb{C})$ has a *polar decomposition* $A = UP$, where $U \in M_n(\mathbb{C})$ is unitary and $P \in M_n(\mathbb{C})$ is positive semidefinite. Thus, every matrix is a product of two normal matrices.

Among normal matrices, positive definite matrices or positive semidefinite matrices are important. Any product of two positive definite (semidefinite) matrices is diagonalizable and has positive (nonnegative) eigenvalues so such products do not fill up $M_n(\mathbb{C})$ or $\mathrm{GL}(n, \mathbb{C})$, the group of $n \times n$ nonsingular matrices. Ballantine [1, 2, 3, 4] showed that every matrix with positive determinant is a product of five positive definite matrices. Radjavi [34] showed that every matrix A with real determinant is a product of at most four Hermitian matrices, and there are matrices which are not product of less. Taussky [37] showed that a matrix A is a product of two positive semidefinite matrices if and only if A is diagonalizable and has nonnegative eigenvalues. The result is extended in the framework of semisimple Lie group by Kostant [29, Proposition 6.2]. Wu [39] extended Taussky's study on positive semidefinite matrices and showed that every matrix with nonnegative determinant is a product of five positive semidefinite matrices. In the same paper, Wu also gave a characterization of matrices that can be expressed as a product of four positive semidefinite matrices. A recent work of Cui, Li, and Sze [9] gives a characterization of products of three positive semidefinite matrices.

In the infinite dimensional setting, Wu [38] showed that on a complex, separable Hilbert space, every unitary operator is a product of sixteen positive operators. Moreover, in the same paper, Wu showed that an operator is a product of finitely many positive operators [1] if and only if it is injective with dense range. In this case, seventeen factors suffice. Phillips [33] improved this and showed that every invertible Hilbert space operator is a product of seven positive operators. Murphy [32] gave a simple proof of Wu's result but without improvements on the number of factors. We also mention Botha's paper [5] which provided a unified treatment of some matrix factorization results, including Ballantine's Theorem. We are interested in the following problem.

Problem. Let G be a subgroup of $\mathrm{GL}(n,\mathbb{C})$ consisting of matrices with positive determinant. Can the matrices in G be written as a product of a finite number of positive definite matrices in G? If so, how many factors do we need to express all elements of G as products of positive definite matrices in G?

We consider the complex symplectic group $\mathrm{Sp}(2n,\mathbb{C})$:

$$\mathrm{Sp}(2n,\mathbb{C}) = \left\{A \in \mathrm{GL}(2n,\mathbb{C}) : A^\top J_n A = J_n\right\},$$

where

$$J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

The symplectic group is a classical group defined as the set of linear transformations of a $2n$-dimensional vector space over $\mathbb{C}$ which preserve the non-degenerate skew-symmetric bilinear

[1]In operator theory, a *positive operator* is what is usually known in matrix theory as a *positive semidefinite* operator.

form which is defined by J_n. It is a non-compact, simply connected, and simple Lie group [21]. Every symplectic matrix $A \in \mathrm{Sp}(2n, \mathbb{C})$ is nonsingular with the inverse $A^{-1} = J_n^{-1} A^\top J_n$. It is obvious to see that $\det A = \pm 1$ but less obvious to narrow it down to $\det A = 1$ though it is true. So one may write

$$\mathrm{Sp}(2n, \mathbb{C}) = \{A \in \mathrm{SL}(2n, \mathbb{C}) : J_n^{-1} A^\top J_n = A^{-1}\}.$$

One can see that $\mathrm{Sp}(2n, \mathbb{C})$ is invariant under $*$, the conjugate transpose. We drop the subscript if the size of the matrix J_n is clear from context. For $A \in \mathrm{GL}(2n, \mathbb{C})$, we define $A^J = J^{-1} A^\top J$, the *J-adjoint* of A. In block form, one can write

$$\begin{aligned} &\mathrm{Sp}(n, \mathbb{C}) \\ = \ &\left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathrm{SL}(2n, \mathbb{C}) : A^\top C = C^\top A, B^\top D = D^\top B, A^\top D - C^\top B = I_n \right\} \end{aligned} \tag{1.1}$$

The following is our main result.

Theorem 1.1. *Each $A \in \mathrm{Sp}(2n, \mathbb{C})$ is a product of five positive definite symplectic matrices.*

Remark: One may be tempted to apply Ballantine's result to get a quick proof. It is true that $A \in \mathrm{Sp}(2n, \mathbb{C})$ can be viewed as a matrix in $\mathrm{GL}(2n, \mathbb{C})$ so by Ballantine's Theorem such A is a product of five positive definite matrices. However, the challenge is that Ballantine Theorem does not assert that those matrices are symplectic as well.

As an example, the symplectic matrix $-I_2$ can be written as the product $-I_2 = P(-P^{-1})$, where

$$P = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}.$$

The matrix P is positive definite with eigenvalues $1, 3$. By [4, Theorem 4], $-P^{-1}$ is a product of 4 positive definite matrices. One can easily check that P is not symplectic, so $-I_2$ can be decomposed into a product of five positive definite matrices and the decomposition is not unique, but at least one is not symplectic. See Corollary 2.2.

Chapter 2

Preliminaries

We start with a discussion on the mathematical tools and theorems that lead to a proof of Theorem 1.1. Most of the material in this chapter can be found in the texts [22, 24, 25]. We also include a section on key results in Ballantine's papers [1, 2, 3, 4] to allow the reader an easy access to the main inspiration for this work. Finally, we end this chapter with a section on symplectic analogues of some key results found in this chapter.

2.1 Matrix equivalence relations and canonical forms

Given an equivalence relation $\sim$ on matrix pairs $(A, B) \in M_n(\mathbb{C})^2$, a *canonical form (or normal form) for* $\sim$ is a complete set of representatives for $\sim$. For the similarity relation

$$A = X^{-1}BX \text{ for some } X \in \mathrm{GL}(n, \mathbb{C}),$$

the most popular choice for a canonical form is the *Jordan canonical form* (JCF). The $k \times k$ Jordan block corresponding to $\lambda \in \mathbb{C}$ is the $k \times k$ upper triangular matrix

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}.$$

A *Jordan matrix* is a direct sum of Jordan blocks.

Theorem 2.1. *Every matrix $A \in M_n(\mathbb{C})$ is similar to a Jordan matrix, which is unique up to permutation of its Jordan blocks.*

Let $A, B \in M_n(\mathbb{C})$ be given. We say that A is *congruent to B if there exists $X \in \mathrm{GL}(n, \mathbb{C})$ satisfying the matrix equation $A = X^*BX$. Another term for this equivalence relation that is used in the literature is *conjunctivity/conjunctive.* If A represents a sesquilinear form $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^n$ under a basis $\mathcal{B}$, then X^*AX is the matrix representation of $\langle \cdot, \cdot \rangle$ under a change of basis $\mathcal{B}$ to $\mathcal{B}'$ given by X. For $A \in \mathrm{GL}(n, \mathbb{C})$, we define its *cosquare to be the matrix $A^{-*}A$. The *cosquare of a matrix A has a particular Jordan Canonical Form [28]:

$$\bigoplus_{j=1}^{p} (J_{m_j}(\mu_j) \oplus J_{m_j}(\overline{\mu_j}^{-1})) \oplus \bigoplus_{k=1}^{q} J_{n_k}(e^{i\theta_k}),$$

where $|\mu_j| > 1$ and $\theta_k \in [0, 2\pi)$. Two matrices that are *congruent have similar *cosquares. However, the converse is not true. As an example, I and $-I$ have the same cosquare but they are not *congruent. If I and $-I$ were *congruent, there would exist $X \in \mathrm{GL}(n, \mathbb{C})$ such that $I = -X^*X$, which makes I negative definite, a contradiction. A "partial" converse that is true is given in [28, Lemma 3.1], which we state below.

Lemma 2.1. (Horn and Sergeichuk, 2006) Let $A, B \in \mathrm{GL}(n, \mathbb{C})$ be given such that $A^{-*}A$ is similar to $B^{-*}B$. Suppose $A^{-*}A = S^{-1}(B^{-*}B)S$ for some $S \in \mathrm{GL}(n, \mathbb{C})$. Set $B_S = S^*BS$ and $M = B_SA^{-1}$. Suppose M has k real eigenvalues counting multiplicities. Then

1. M is similar to a real matrix.

2. There exists $D_- \in M_k(\mathbb{C})$ and $D_+ \in M_{n-k}(\mathbb{C})$ such that A is *congruent to $-D_- \oplus D_+$ and B is *congruent to $D_- \oplus D_+$.

For every natural number $n \in \mathbb{N}$, we define the matrices

$$H_{2n}(\mu) = \begin{bmatrix} 0 & I_n \\ J_n(\mu) & 0 \end{bmatrix}$$

and

$$D_n = \begin{bmatrix} & & & & (-1)^{n+1} \\ & & & \iddots & (-1)^n \\ & -1 & -1 & \iddots & \\ 1 & 1 & & & \end{bmatrix},$$

with $D_1 = [1]$. Note that $H_{2n}(\mu)^{-*}H_{2n}(\mu)$ is similar to $J_n(\mu) \oplus J_n(\overline{\mu}^{-1})$ and $D_n^{-*}D_n$ is similar to $J_n((-1)^{n+1})$. The matrices $H_{2n}(\mu)$ and D_n together with the Jordan blocks $J_n(0)$ are the canonical blocks of Horn and Sergeichuk's canonical form for *congruence [28, Theorem 1.1].

Theorem 2.2. (Horn and Sergeichuk, 2006) Each square complex matrix is *congruent to a direct sum, uniquely determined up to permutation of summands, of canonical matrices of the following three types:

1. $J_n(0)$,
2. λD_n, where $|\lambda| = 1$, and
3. $H_{2n}(\mu)$, where $|\mu| > 1$.

Instead of D_n, one may use any other nonsingular matrix F_n for which there exists a real θ_n such that $F_n^{-*}F_n$ is similar to $J_n(e^{i\theta_n})$.

2.2 Tridiagonal pseudo-Toeplitz matrices

A *tridiagonal matrix* is a matrix whose entries above the first superdiagonal and below the first subdiagonal are all zeros. The matrix

$$\begin{bmatrix} 2 & 3 & 0 & 0 & 0 \\ 3 & 1 & 4 & 0 & 0 \\ 0 & -6 & 1 & 3 & 0 \\ 0 & 0 & 8 & 8 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

is an example of a tridiagonal matrix. A *Toeplitz matrix* is a matrix whose entries on a diagonal are equal, for example,

$$\begin{bmatrix} 1 & 3 & 2 \\ 4 & 1 & 3 \\ -2 & 4 & 1 \end{bmatrix}$$

is a Toeplitz matrix. We use the notation $T_k(a, b, c)$ for a tridiagonal Toeplitz matrix with main diagonal entries a, first subdiagonal entries b, and first superdiagonal entries c. Let $\phi_k(a)(\lambda)$ be the characteristic polynomial of $T_k(a, 1, 1)$. Expanding $\det(T_k(a, 1, 1) - \lambda I)$ by the last row reveals the following recursive identity:

$$\phi_k(a)(\lambda) = (a - \lambda)\phi_{k-1}(a)(\lambda) - \phi_{k-2}(a)(\lambda)$$

for $k \geq 2$ with $\phi_0(a) = 1$ and $\phi_1(a) = a - \lambda$. Using the substitution $a - \lambda = 2x$, we get that $\phi_k(a)(x) = U_k(x)$, the kth degree Chebyshev polynomial of the second kind [30, p.65]. Moreover, if $\phi_k(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc})$ is the characteristic polynomial of

$$T_k(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc}),$$

then we have $\phi_k(a/\sqrt{bc}, b/\sqrt{bc}, c/\sqrt{bc}) = \phi_k(a/\sqrt{bc}) = U_k$ [30, p.66].

The matrix D_n in Horn and Sergeichuk's canonical form has an unusual structure. Hence, essential properties of D_n, such as information on its eigenvalues, are not immediately obtained by inspection. Observe that

$$D_k^2 = (-1)^k \begin{bmatrix} -1 & -1 & & & & \\ 1 & 0 & -1 & & & \\ & 1 & 0 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & -1 \\ & & & & 1 & 0 \end{bmatrix},$$

which we could write as a sum of a tridiagonal Toeplitz matrix and the standard basis matrix E_{11}:

$$D_k^2 = (-1)^k [T_k(0,1,-1) - E_{11}].$$

Let ψ_k be the characteristic polynomial of $C_k = -i[T_k(0,1,-1) - E_{11}]$. We obtain that

$$\psi_k(\lambda) = \phi_k(0,-i,i)(\lambda) + i\phi_{k-1}(0,-i,i)(\lambda) = U_k\left(-\frac{\lambda}{2}\right) + iU_{k-1}\left(-\frac{\lambda}{2}\right).$$

We show that ψ_k has simple roots. To do this, it suffices to show that the *discriminant* $D(\psi_k)$ of ψ_k is nonzero. The discriminant can be defined in terms of the following matrix [6, Definition 7]. Given polynomials $f, g \in \mathbb{C}[x]$ of positive degree, write them in the form

$$f = a_0 x^l + \cdots + a_l, a_0 \neq 0,$$

$$g = b_0 x^m + \cdots + b_m, b_0 \neq 0.$$

The *Sylvester matrix* of f and g with respect to x is

$$S(f,g,x) = \begin{bmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ a_2 & a_1 & \ddots & & b_2 & b_1 & \ddots & \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ & \vdots & & a_1 & & \vdots & & b_1 \\ a_l & & & & b_m & & & \\ & a_l & & \vdots & & b_m & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_l & & & & b_m \end{bmatrix} \in M_{m+l}(\mathbb{C}).$$

The discriminant of f with respect to x is

$$D(f) = \frac{(-1)^{l(l-1)/2}}{a_0} \det(S(f,f',x)),$$

where f' is the derivative of f with respect to x. It is known that f has multiple roots if and only if $D(f) = 0$ [6, Exercise 5.8]. The following theorem by Dilcher and Stolarsky [10, Theorem 4] provides an explicit formula for the discriminant of $U_n + kU_{n-1}$.

Theorem 2.3. (Dilcher and Stolarsky, 2004) For all $n \geq 1$, we have

$$D(U_n + kU_{n-1}) = 2^{n(n-1)} a_{n-1}(k),$$

where

$$a_{n-1}(k) = (-1)^n \frac{(2n+1)^n k^n}{(n+1)^2 - n^2 k^2} \left[U_n \left(-\frac{n+1+nk^2}{(2n+1)k} \right) + kU_{n-1} \left(-\frac{n+1+nk^2}{(2n+1)k} \right) \right]$$

is an even polynomial in k of degree $2n-2$ with positive integer coefficients.

Our characteristic polynomial correspond to the case $k = i$:

$$a_{n-1}(i) = \frac{(-1)^n (2n+1)^n i^n}{(n+1)^2 + n^2} \left[U_n \left(\frac{i}{2n+1} \right) + iU_{n-1} \left(\frac{i}{2n+1} \right) \right].$$

An explicit expression for $U_n(x)$ is given in [41, Section 6.10.7]:

$$U_n(x) = \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \begin{pmatrix} n-m \\ m \end{pmatrix} (2x)^{n-2m}.$$

Observe that

$$\left(\frac{2i}{2n+1} \right)^{n-2m} = \left(\frac{2}{2n+1} \right)^{n-2m} i^{n-2m} = (-1)^m i^n \left(\frac{2}{2n+1} \right)^{n-2m}.$$

Therefore,

$$U_n \left(\frac{i}{2n+1} \right) = i^n \sum_{m=0}^{\lfloor n/2 \rfloor} \begin{pmatrix} n-m \\ m \end{pmatrix} \left(\frac{2}{2n+1} \right)^{n-2m},$$

where the sum is positive. It follows that

$$U_n\left(\frac{i}{2n+1}\right)+iU_{n-1}\left(\frac{i}{2n+1}\right)$$

is of the form $i^n M$ for some $M>0$. This shows that $a_{n-1}(i)\neq 0$. It follows that ψ_k has simple roots. Therefore, C_k has distinct eigenvalues. Consequently, D_k^2 has distinct eigenvalues as well. Since D_k is a square root of D_k^2, D_k also has distinct eigenvalues.

Proposition 2.1. *The matrix D_k has distinct eigenvalues for all $k\in\mathbb{N}$.*

2.3 Matrix functions

Let f be a complex-valued scalar function. The matrix function $f(A)$, where $A\in M_n(\mathbb{C})$, is generally defined using any of the following three equivalent definitions [22, Chapter 1].

1. **Jordan Canonical Form.** Let f be defined on $\sigma(A)\subseteq\mathbb{C}$ and let $A=X^{-1}f(J)X$, where $J=\bigoplus_{j=1}^m J_{k_j}(\lambda_j)$ is a Jordan matrix. Then

$$f(A)=X^{-1}\bigoplus_{j=1}^{m} f(J_{k_j}(\lambda_j))X,$$

where

$$f(J_{k_j}(\lambda_j))=\begin{bmatrix} f(\lambda_j) & f'(\lambda_j) & \cdots & \frac{f^{(k_j-1)}(\lambda_j)}{(k_j-1)!} \\ & f(\lambda_j) & \ddots & \vdots \\ & & \ddots & f'(\lambda_j) \\ & & & f(\lambda_j) \end{bmatrix}$$

for all $j=1,\ldots,m$.

2. **Polynomial Interpolation.** Let f be defined on $\sigma(A)\subseteq\mathbb{C}$ and $\psi(\lambda)=\prod_{j=1}^{s}(\lambda-\lambda_j)^{n_j}$ be the minimal polynomial of A. Then

$$f(A)=p(A),$$

where p is the *Hermite interpolating polynomial* for ψ. That is, p is the unique polynomial with $\deg p < \deg \psi$ and satisfies

$$p^{(j)}(\lambda_k) = f^{(j)}(\lambda_k)$$

for all $j = 0, \dots, n_k - 1$, and $k = 1, \dots, s$.

3. **Cauchy Integral.** Let f be analytic on and inside a closed contour Γ that encloses $\sigma(A)$. Then
$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} \, dz.$$

For the proof of the equivalence of these statements, see [22, Theorem 1.12]. An immediate application of matrix functions in the study of factorizations is seen in the following. A *complex orthogonal* matrix $Q \in M_n(\mathbb{C})$ is a matrix satisfying $Q^\top Q = I$.

Theorem 2.4. [25, Corollaries 6.4.18 and 6.4.19] *Let $A, B \in M_n(\mathbb{C})$ be given and suppose there is a single polynomial $p(t)$ such that $A^\top = p(A)$ and $B^\top = p(B)$. Then A and B are similar if and only if they are similar via a complex orthogonal matrix. In particular, this is true if A and B are both (a) symmetric, (b) skew-symmetric, (c) complex orthogonal.*

Proof. If $A, B \in M_n(\mathbb{C})$ are similar, then there exists $X \in GL(n, \mathbb{C})$ such that $A = X^{-1}BX$. For any polynomial $p(t)$, we have $p(A) = X^{-1}p(B)X$. By hypothesis, it follows that $A^\top = X^{-1}B^\top X$. Taking the transpose of both sides, we get $A = X^\top B X^{-1^\top}$. That is, $B(XX^\top) = (XX^\top)B$. In particular, any polynomial in $XX^\top$ commutes with B. Let S be a symmetric square root of $XX^\top$ obtained by interpolating the primary square root function $r(z) = \sqrt{z}$ on the eigenvalues of $XX^\top$. If g is the Hermite interpolating polynomial, then $S = g(XX^\top)$ satisfies $S^2 = XX^\top$. Moreover, $BS = SB$. Set $Q = S^{-1}X$. That is, $X = SQ$. Thus,

$$A = (SQ)^{-1}B(SQ) = Q^{-1}S^{-1}(BS)Q = Q^{-1}S^{-1}(SB)Q = Q^{-1}BQ.$$

To verify that Q is complex orthogonal, we compute

$$Q^\top Q = (S^{-1}X)^\top (S^{-1}X) = X^\top S^{-2} X = X^\top (XX^\top)^{-1} X = I.$$

If $A^\top = A$, we have $p(t) = t$; if $A^\top = -A$, then $p(t) = -t$; if $A^\top = A^{-1}$, we take $p(t)$ to be the Hermite interpolating polynomial of $f(t) = \dfrac{1}{t}$ on $\sigma(A) = \sigma(B)$. $\square$

The following theorem [23, Theorem 3.2] characterizes matrix functions $f(\overline{A}) = \overline{f(A)}$. We present the proof that was given in [23].

Theorem 2.5. *Let f be analytic on an open subset $\Omega \subseteq \mathbb{C}$ such that each connected component of Ω is closed under conjugation. Consider the corresponding matrix function f on its natural domain in $M_n(\mathbb{C})$, the set $\mathcal{D} = \{A \in M_n(\mathbb{C}) : \sigma(A) \subseteq \Omega\}$. The following are equivalent.*

1. $f(\overline{A}) = \overline{f(A)}$ *for all* $A \in \mathcal{D}$.
2. $f(\mathcal{D} \cap M_n(\mathbb{R})) \subseteq M_n(\mathbb{R})$.
3. $f(\mathbb{R} \cap \Omega) \subseteq \mathbb{R}$.

Proof. We do a round robin proof. Suppose (1) holds. If $A \in M_n(\mathbb{R})$, then $A = \overline{A}$. It follows that $f(A) = f(\overline{A}) = \overline{f(A)}$, that is, $f(A) \in M_n(\mathbb{R})$. Therefore, (1) implies (2). Suppose (2) holds. If $\lambda \in \mathbb{R} \cap \Omega$, then $\lambda I \in \mathcal{D}$. By (2), $f(\lambda I) \in M_n(\mathbb{R})$. Since $f(\lambda I) = f(\lambda) I$, we have $f(\lambda) \in \mathbb{R}$. Therefore, (2) implies (3). Finally, we show (3) implies (1). Let $\tilde{\Omega}$ be a connected component of Ω. Since $\tilde{\Omega}$ is open and connected, it is path-connected, and since it is closed under conjugation it must contain some $\lambda \in \mathbb{R}$ by the intermediate value theorem. Since $\tilde{\Omega}$ is open, the set $U = \tilde{\Omega} \cap \mathbb{R}$ is a nonempty open subset of $\mathbb{R}$. By hypothesis, $f(U) \subseteq \mathbb{R}$. Define the functions $g(z) = f(\overline{z})$ and $h(z) = \overline{f(z)}$. Then $g = h$ on the set U. Since U contains a limit point, by the identity theorem for holomorphic functions, $g = h$ on $\tilde{\Omega}$. This argument holds for all the other connected components of Ω, so $f(\overline{z}) = \overline{f(z)}$ on Ω. It follows that $f(\overline{A}) = \overline{f(A)}$ holds for all diagonal matrices in $\mathcal{D}$, and consequently, for all diagonalizable matrices in $\mathcal{D}$. Since the scalar function f is analytic on Ω, the matrix function f is continuous on $\mathcal{D}$. The diagonalizable matrices form a dense open subset of $\mathcal{D}$. Therefore, the statement (1) holds for all matrices in $\mathcal{D}$. $\square$

2.4 Ballantine's theory.

Ballantine's characterization of products of positive definite matrices relies on the following result, presented in a more general algebraic setting. Let G be a group with identity element e. For nonempty subsets S, T of G, we define the usual set product ST to be

$$ST = \{st : s \in S, t \in T\}$$

and the set power T^k to be

$$T^k = \{t_1 t_2 \cdots t_k : t_1, \ldots, t_k \in T\}$$

for any $k \in \mathbb{N}$, which denotes the set of natural numbers. We define $T^0 = \{e\}$. The main characterization of products of positive definite matrices in any group can be obtained from the following result due to Ballantine [3].

Theorem 2.6. (Ballantine [3, Theorem 1]) *Let G be a group with identity element e and $S \subseteq G$. Let $\star$ be an involutory anti-automorphism on G. Define the sets*

$$\phi(T) = \bigcup_{x \in G} x^{-1} T x$$

and

$$\psi(T) = \bigcup_{x \in G} x^{\star} T x$$

for any $T \subseteq G$. Let $F = \psi(S)$. For any $m \in \mathbb{N} \cup \{0\}$, we have

1. *$\phi(F^{2m}) = F^{2m} = \phi(SF^{2m-1})$, and*

2. *$\psi(F^{2m+1}) = F^{2m+1} = \psi(SF^{2m})$.*

Here, $\phi(SF^{-1})$ is defined to be the trivial subgroup $\{e\}$.

Take $G = \mathrm{GL}(n, \mathbb{C})$, with the conjugate transpose map $* : A \mapsto A^*$ and $S = \{I\}$. Then $F = \psi(S)$ is the set of all matrices that are *congruent to I. Clearly these matrices are the positive definite matrices. The set F^k is the set of all products of k positive definite matrices; the set $\phi(F^k)$ is the set of all matrices which are similar to a matrix in F^k; the set $\psi(F^k)$ is the set of all matrices which are *congruent to a matrix in F^k. This choice of G, $*$, and S gives us the basis for Ballantine's characterizations of products of positive definite matrices [4, Theorem 1].

Theorem 2.7. *Let $A \in GL(n, \mathbb{C})$ be given and let $k \in \mathbb{N}$. Then statements $(1) - (3)$ are equivalent and statements $(4) - (6)$ are equivalent.*

1. *A is a product of $2k$ positive definite matrices.*

2. *A is similar to a product of $2k$ positive definite matrices.*

3. *A is similar to a product of $2k - 1$ positive definite matrices.*

4. *A is a product of $2k + 1$ positive definite matrices.*

5. *A is *congruent to a product of $2k + 1$ positive definite matrices.*

6. *A is *congruent to a product of $2k$ positive definite matrices.*

Next, we present Ballantine's further characterizations for products of a fixed number of positive definite matrices. We start with the most simple case, that of products of two positive definite matrices. We have the following [4, Theorem 2].

Theorem 2.8. *Let $A \in M_n(\mathbb{C})$. The following are equivalent.*

1. *A is a product of 2 positive definite matrices.*

2. *A is similar to a product of 2 positive definite matrices.*

3. *A is similar to a positive definite matrix*

4. *A is unitarily similar to a diagonalizable lower triangular matrix with positive diagonal entries.*

Proof. The equivalence of (1) - (3) is due to Theorem 2.7. Suppose A is similar to a positive definite matrix P. By Schur's Triangularization Theorem [24, p.101-102], $A = U^*LU$, for some $U \in \mathrm{U}(n)$ and lower triangular $L \in M_n(\mathbb{C})$. Since A is similar to P, then $\sigma(A) \subseteq (0, \infty)$. Thus, the diagonal entries of L are positive. Note that L and P are similar and P is diagonalizable, so L is diagonalizable as well. For the converse, suppose $A = W^*TW$, where $W \in U(n)$ and T is a diagonalizable lower triangular matrix with positive diagonal entries. Since the diagonal entries of T are its eigenvalues, T is similar to a diagonal matrix P with positive eigenvalues. It follows that A is also similar to the positive definite diagonal matrix P. □

Sourour [36] gave the conditions on how one can choose the eigenvalues of matrices B and C in a factorization $A = BC$. The precise statement is given in the theorem below.

Theorem 2.9 (Sourour [36]). *Let $\mathbb{F}$ be a field and $A \in M_n(\mathbb{F})$ be a nonsingular nonscalar matrix. Let $\mathbf{b} = (b_1, \dots, b_n), \mathbf{c} = (c_1, \dots, c_n) \in \mathbb{F}^n$ such that*

$$\prod_{j=1}^{n} b_j c_j = \det A.$$

Then there exist matrices $B, C \in M_n(\mathbb{F})$ such that $A = BC$, where $(b_1, \dots, b_n)$, and $(c_1, \dots, c_n)$ are the vectors of eigenvalues of B and C, respectively.

Proof. We use induction to prove the theorem. The result is vacuously true for the case $n = 1$. Suppose the conclusion holds for all matrices of size less than n. Let $A \in M_n(\mathbb{C})$. First, we show that A is similar to a matrix whose $(1, 1)$ entry is b_1c_1. Since A is nonscalar, $A - b_1c_1I$ is nonscalar and there exists a nonzero $x_1 \in \mathbb{C}^n$ such that $(A - b_1c_1I)x_1 \neq 0$ and x_1 is not an eigenvector of $A - b_1c_1I$. Let $x_2 = (A - b_1c_1I)x_1$. Since x_1 is not an eigenvector of $A - b_1c_1I$, x_1 and x_2 are linearly independent. Extend the set $\{x_1, x_2\}$ into an ordered basis $\mathcal{B} = \{x_1, x_2, \dots, x_n\}$ of $\mathbb{C}^n$. If $\tilde{A}$ is the linear transformation given by $\tilde{A} : x \mapsto Ax$, then relative to the basis $\mathcal{B}$, the matrix representation A_1 of $\tilde{A}$ will have first column

$$\begin{bmatrix} b_1c_1 & 1 & 0 & \cdots & 0 \end{bmatrix}^\top.$$

We conclude that A is similar to the matrix A_1, which we can partition as

$$A_1 = \begin{bmatrix} b_1c_1 & y^\top \\ x & R \end{bmatrix},$$

where $x \neq 0$. If $n = 2$, then

$$\begin{bmatrix} b_1c_1 & y \\ x & r \end{bmatrix} = \begin{bmatrix} b_1 & 0 \\ c_1^{-1}x & b_2 \end{bmatrix} \begin{bmatrix} c_1 & b_1^{-1}y \\ 0 & c_2 \end{bmatrix}.$$

Equality holds since the $(2, 2)$-entry of the right-hand side is equal to

$$\begin{aligned} b_1^{-1}c_1^{-1}xy + b_2c_2 &= b_1^{-1}c_1^{-1}(xy + \det A) \\ &= b_1^{-1}c_1^{-1}(xy + \det A_1) \\ &= b_1^{-1}c_1^{-1}(xy + b_1c_1r - xy) \\ &= r \end{aligned}$$

and the other entries are easily checked to be the equal on both sides. Thus, the conclusion of the theorem holds for $n = 2$. Suppose $n \geq 3$. We show that A_1 is similar to a matrix of the form

$$A_2 = \begin{bmatrix} b_1c_1 & z^\top \\ x & S \end{bmatrix},$$

where $S - b_1^{-1}c_1^{-1}xz^\top$ is a nonscalar matrix. If $R - b_1^{-1}c_1^{-1}xy^\top$ is nonscalar, we are done. Suppose $R - b_1^{-1}c_1^{-1}xy^\top = \alpha I$ for some $\alpha \in \mathbb{C}$. Since rank $A > 2$, the span of the columns of R is not contained in span $\{x\}$. This implies that there exists $w \in C^{n-1}$ such that $w^\top x = 0$ but

$w^\top R \neq 0$. Let $P = \begin{bmatrix} 1 & w^\top \\ 0 & I \end{bmatrix}$. Then

$$\begin{aligned} P^{-1}A_1P &= \begin{bmatrix} 1 & -w^\top \\ 0 & I \end{bmatrix} \begin{bmatrix} b_1c_1 & y^\top \\ x & R \end{bmatrix} \begin{bmatrix} 1 & w^\top \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} b_1c_1 & y^\top + b_1c_1w^\top - w^\top R \\ x & R + xw^\top \end{bmatrix}. \end{aligned}$$

Let $S = R + xw^\top$ and $z^\top = y^\top + b_1c_1w^\top - w^\top R$. Then

$$\begin{aligned} S - b_1^{-1}c_1^{-1}xz^\top &= R + xw^\top - b_1^{-1}c_1^{-1}xz^\top \\ &= R + xw^\top - b_1^{-1}c_1^{-1}x(y^\top + b_1c_1w^\top - w^\top R) \\ &= (R - b_1^{-1}c_1^{-1}xy^\top) - b_1^{-1}c_1^{-1}xw^\top R \\ &= \alpha I - b_1^{-1}c_1^{-1}xw^\top R. \end{aligned}$$

Since both $x \neq 0$ and $w^\top R \neq 0$, we have $\operatorname{rank}(b_1^{-1}c_1^{-1}xw^\top R) = 1$. We conclude that $S - b_1^{-1}c_1^{-1}xz^\top$ is nonscalar. Note that

$$\begin{bmatrix} b_1c_1 & z^\top \\ x & S \end{bmatrix} = \begin{bmatrix} b_1c_1 & 0 \\ x & S - b_1^{-1}c_1^{-1}xz^\top \end{bmatrix} \begin{bmatrix} 1 & b_1^{-1}c_1^{-1}z^\top \\ 0 & I \end{bmatrix}$$

Thus,

$$\det(S - b_1^{-1}c_1^{-1}xz^\top) = b_1^{-1}c_1^{-1}\det A = \prod_{j=2}^{n} b_jc_j.$$

By induction hypothesis, there exist matrices $B', C' \in M_n(\mathbb{C})$ such that

$$\sigma(B') = \{b_2, \cdots, b_n\}, \quad \sigma(C') = \{c_2, \ldots, c_n\},$$

and $S - b_1^{-1}c_1^{-1}xz^\top = B'C'$. It follows that

$$A_2 = P^{-1}A_1P = \begin{bmatrix} b_1 & 0 \\ c_1^{-1}x & B' \end{bmatrix} \begin{bmatrix} c_1 & b_1^{-1}z^\top \\ 0 & C' \end{bmatrix}.$$

This proves the theorem. $\square$

Let $A \in M_n(\mathbb{C})$ be given such that A is nonscalar and $\det A > 0$. By Theorem 2.9, $A = BC$, where we can choose B and C such that both B and C have distinct positive eigenvalues. Thus, both B and C are similar to positive definite diagonal matrices. By Theorem 2.8, there exists positive definite matrices $P_1, P_2, P_3, P_4 \in M_n(\mathbb{C})$ such that $B = P_1P_2$ and $C = P_3P_4$. Thus, $A = P_1P_2P_3P_4$. Let $\lambda \in \mathbb{C}$. Suppose the matrix λI is a product of four positive definite matrices. We have $\lambda I = BC$, where B and C have positive eigenvalues. Thus, $\lambda B^{-1} = C$ and for each $c \in \sigma(C)$, there exists $b \in \sigma(B)$ such that $c = \lambda b^{-1}$. Since $b, c > 0$, it follows that $\lambda > 0$. That is, the only scalar matrices that are products of four (or less) positive definite matrices are positive scalar matrices.

Theorem 2.10. *Let $A \in M_n(\mathbb{C})$. The following are equivalent.*

1. *A is a product of 4 positive definite matrices.*

2. *A is similar to a product of 4 positive definite matrices.*

3. *A is similar to a product of 3 positive definite matrices.*

4. *A is nonscalar with $\det A > 0$ or $A = \lambda I$ for some $\lambda > 0$.*

Consider $A = \lambda I \in M_n(\mathbb{C})$, where $\det A = \lambda^n > 0$. Let $P \in M_n(\mathbb{C})$ be a nonscalar positive definite matrix. Then $A = P(\lambda P^{-1})$. By the previous theorem, λP^{-1} is a product of four positive definite matrices. Hence, A is a product of five positive definite matrices.

Theorem 2.11. *Let $A \in M_n(\mathbb{C})$ be given such that $\det A > 0$. Then A is a product of five positive definite matrices.*

2.5 Symplectic analogues

Now, let us take our group G to be the symplectic group $\text{Sp}(2n, \mathbb{C})$, the involutory anti-automorphism $\star$ to be the conjugate transpose map $*$, and the set $S = \{I\}$. Then

$$\psi(S) = \{X^* X : X \in \text{Sp}(2n, \mathbb{C})\}$$

is the set of all positive definite symplectic matrices. Indeed, the product $X^* X$ is positive definite symplectic. Now, for the converse, let $A \in \text{Sp}(2n, \mathbb{C})$ be positive definite. Its unique positive definite square root $A^{1/2}$ is also symplectic [31, Theorem 3.2]. Thus, $A = (A^{1/2})^* A^{1/2}$. Another way to see that $A^{1/2}$ is symplectic is from Lie Theory [21]. The Lie algebra of $\text{Sp}(2n, \mathbb{C})$ is

$$\mathfrak{sp}(2n, \mathbb{C}) = \{X \in M_{2n}(\mathbb{C}) : J_n X + X^\top J_n = 0\}.$$

and it admits a Cartan decomposition θ such that $\mathfrak{sp}(2n, \mathbb{C}) = \mathfrak{k} + \mathfrak{p}$ in which $\mathfrak{p}$ is the -1-eigenspace and $\mathfrak{k}$ is the $+1$-eigenspace of θ and this θ can be lifted to the global Cartan decomposition $\Theta : G \to G$ [21] such that

1. K, the subgroup of G having $\mathfrak{k}$ as the Lie algebra, is the group invariant under G.

2. the map $K \times \mathfrak{p} \to G$ such that $(k, X) \mapsto ke^K$.

So $\exp : \mathfrak{p} \to P = \text{Sp}(2n, \mathbb{C}) \cap \mathbb{P}_n$ is a diffeomorphism. Thus, given $A \in \text{Sp}(2n, \mathbb{C}) \cap \mathbb{P}_n$ there is $X \in \mathfrak{p}$ such that $\exp X = A$ and $A^{1/2} = \exp(X/2)$.

Remark: Since $\text{Sp}(2n, \mathbb{C})$ is invariant under $*$ and the square root of positive definite symplectic matrix is also positive definite symplectic, we conclude that [35, p.188] the polar decomposition for $\text{Sp}(2n, \mathbb{C})$ comes from the usual polar decomposition of $\text{GL}(2n, \mathbb{C})$:

$$\text{GL}(2n, \mathbb{C}) = \text{U}(2n)\mathbb{P}_{2n}$$

where $\mathrm{U}(2n)$ is the unitary group of order $2n$ and $\mathbb{P}_{2n}$ is the space of $2n \times 2n$ positive definite matrices. In other words,

$$\mathrm{Sp}(2n, \mathbb{C}) = KP, \tag{2.1}$$

where $P = \mathbb{P}_{2n} \cap \mathrm{Sp}(2n, \mathbb{C})$ and $K = \mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{C})$.

Two matrices $A, B \in M_{2n}(\mathbb{C})$ are said to be *symplectically similar* if there is a symplectic matrix of similarity between A and B; A and B are said to be *symplectically* **congruent* if there is a symplectic matrix of *congruence between A and B, i.e. $A = X^*BX$ for some $X \in \mathrm{Sp}(2n, \mathbb{C})$. From our chosen G, $\star$, and S, we get the following.

Corollary 2.1. *Let* $A \in \mathrm{Sp}(2n, \mathbb{C})$ *be given and let* $k \in \mathbb{N}$. *Then statements* $(1) - (3)$ *are equivalent and statements* $(4) - (6)$ *are equivalent.*

1. *A is a product of $2k$ positive definite symplectic matrices.*

2. *A is symplectically similar to a product of $2k$ positive definite symplectic matrices.*

3. *A is symplectically similar to a product of $2k - 1$ positive definite symplectic matrices.*

4. *A is a product of $2k + 1$ positive definite symplectic matrices.*

5. *A is symplectically *congruent to a product of $2k+1$ positive definite symplectic matrices.*

6. *A is symplectically *congruent to a product of $2k$ positive definite symplectic matrices.*

It is known that if $A, B \in \mathrm{Sp}(2n, \mathbb{C})$ are similar, then the matrix of similarity can be taken to be symplectic [26, Corollary 22]. Its proof is based on Horn and Johnson's proof that the matrix of similarity of complex orthogonal matrices can be taken to be complex orthogonal as well. We make adjustments to their proof to show that the same holds for *congruence.

Theorem 2.12. *Let* $A, B \in \mathrm{Sp}(2n, \mathbb{C})$ *be given. If A is *congruent to B, then the matrix of *congruence can be taken to be symplectic.*

Proof. Suppose $A = X^*BX$ for some nonsingular $X \in M_n(\mathbb{C})$. Note that $A^J = X^J B^J (X^J)^*$. Since $A^J = A^{-1}$, $B^J = B^{-1}$, taking inverse on both sides gives us $A = (X^J)^{-*}B(X^J)^{-1}$. Thus, we have $BY = Y^{-*}B$, where $Y = XX^J$. Observe that $Y^J = Y$. Since $e^{i\theta}X$ is another matrix of congruence between A and B, without loss of generality, we can pick X such that $\sigma(Y) \subseteq \mathbb{C}\backslash(-\infty, 0]$ (otherwise, choose $\theta \in \mathbb{R}$ such that $e^{2i\theta}Y$ does not have negative eigenvalues). Let p be a polynomial function interpolating the principal square root function f on the union of the spectra $\sigma(Y) \cup \sigma(Y^{-*})$. The principal square root f is analytic on the open connected set $\mathbb{C}\backslash(-\infty, 0]$. This set is closed under conjugation and f is real whenever $t > 0$. By Theorem 2.5, $f(\overline{C}) = \overline{f(C)}$ for all C with $\sigma(C) \subseteq \mathbb{C}\backslash(-\infty, 0]$. In this case, p can also be chosen to be a real polynomial. Hence, we have

$$p(Y)^{-*} = f(Y)^{-*} = f(Y^{-*}) = p(Y^{-*})$$

and

$$p(Y)^{-*}B = p(Y^{-*})B = Bp(Y).$$

Let $Z = p(Y)^{-1}X$. Then

$$Z^J Z = X^J p(Y^J)^{-1} p(Y)^{-1} X = X^J p(Y)^{-2} X = X^J Y^{-1} X = X^J (XX^J)^{-1} X = I.$$

Thus, $Z \in \mathrm{Sp}(2n, \mathbb{C})$ and $X = p(Y)Z$. Finally,

$$A = X^* B X = Z^* p(Y)^* B p(Y) Z = Z^* p(Y)^* p(Y)^{-*} B Z = Z^* B Z,$$

as desired. $\square$

Remark. By the previous theorem, we can replace symplectic similarity/*congruence by general similarity/*congruence in Corollary 2.1.

2.5.1 Products of two positive definite symplectic matrices

We give more characterizations of products of two positive definite symplectic matrices that is analogous to Ballantine's results.

Theorem 2.13. *Let $A \in \mathrm{Sp}(2n, \mathbb{C})$ be given. The following are equivalent.*

1. *A is a product of two positive definite symplectic matrices.*

2. *A is similar to a product of two positive definite symplectic matrices.*

3. *A is similar to a positive definite symplectic matrix.*

4. *A is diagonalizable and $\sigma(A) \subseteq (0, \infty)$.*

5. *A is unitarily similar to a matrix of the form $X \oplus (X^{-1})^\top$, where X is a product of two positive definite matrices.*

6. *A is a product of two positive definite matrices.*

Proof. The equivalence of statements (1), (2), (3), is given by Corollary 2.1. If A is similar to a positive definite symplectic matrix B, then $\sigma(A) \subseteq (0, \infty)$. The matrix B is normal, hence unitarily diagonalizable. It follows that A is diagonalizable. This shows that (3) implies (4). Suppose A is diagonalizable and $\sigma(A) \subseteq (0, \infty)$. Since A is symplectic and the eigenvalues are positive, we have pairs $\lambda, \lambda^{-1} \in \sigma(A)$ and the algebraic multiplicity of 1 as an eigenvalue is even. Thus, a matrix $X \oplus X^\top$ satisfying (5) can be constructed. If X has positive eigenvalues, then X is a product of two positive definite matrices [4, Theorem 2]. □

2.5.2 2×2 Symplectic matrices

A matrix $A \in M_2(\mathbb{C})$ is symplectic if and only if $\det A = 1$. That is, $\mathrm{Sp}(2, \mathbb{C}) = \mathrm{SL}(2, \mathbb{C})$, the special linear group. Let $A \in M_n(\mathbb{C})$ be a nonsingular nonscalar matrix. Recall that $\sigma(X)$ denotes the spectrum of $X \in M_n(\mathbb{C})$. We denote by $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X)) \in \mathbb{C}^n$ the vector of eigenvalues of X (unique up to permutations). A direct consequence of Sourour's theorem is the following.

Corollary 2.2. *Every matrix $A \in \mathrm{Sp}(2, \mathbb{C})$ is a product of five positive definite symplectic matrices. If $A \neq -I$, then A can be expressed as a product of four positive definite symplectic matrices.*

Proof. The only scalar matrices in $\mathrm{Sp}(2, \mathbb{C})$ are I and $-I$. The identity matrix I is a positive definite symplectic matrix, so it is a product of four positive definite symplectic matrices. Let $A \in \mathrm{Sp}(2, \mathbb{C})$. Suppose $A \neq I, -I$. By Sourour's theorem, $A = BC$, where $\lambda(B) = (\beta, \beta^{-1})$, and $\lambda(C) = (\gamma, \gamma^{-1})$. Notice that $B, C \in \mathrm{Sp}(2, \mathbb{C})$. Choose $\beta, \gamma > 1$. By Theorem 2.13, B and C are both products of two positive definite symplectic matrices. Hence, A is a product of four positive definite symplectic matrices. Finally, we can write $-I = P(-P^{-1})$ for some nonscalar positive definite symplectic P. The matrix $-P^{-1}$ is nonscalar so it is a product of four positive definite symplectic matrices. $\square$

2.5.3 Canonical forms

Recall that by Corollary 2.1, *congruence preserves the property of a matrix to be written as a product of five positive definite symplectic matrices. Thus, if we have a canonical form for symplectic *congruence, it suffices to check each canonical block that will arise. In this section, we construct a symplectic *congruence canonical form using the symplectic Jordan canonical form.

Symplectic matrices have a special Jordan Canonical Form. Let $A = [A_{ij}] \in M_{2m}(\mathbb{C})$ and $B = [B_{ij}] \in M_{2n}(\mathbb{C})$, where $A_{ij} \in M_m(\mathbb{C})$ and $B_{ij} \in M_n(\mathbb{C})$ for $i, j \in \{1, 2\}$. The *expanding sum* of A and B is

$$A \boxplus B = \begin{bmatrix} A_{11} \oplus B_{11} & A_{12} \oplus B_{12} \\ A_{21} \oplus A_{21} & A_{22} \oplus B_{22} \end{bmatrix} \in M_{2(m+n)}(\mathbb{C}).$$

We remark that the expanding sum is different from the usual direct sum

$$A \oplus B = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ 0 & 0 & B_{11} & B_{12} \\ 0 & 0 & B_{21} & B_{22} \end{bmatrix} \in M_{2(m+n)}(\mathbb{C}).$$

Note that $A \boxplus B$ is symplectic if and only if A and B are symplectic. Moreover, $A \boxplus B$ is similar to $A \oplus B$. Define the matrix

$$F_{2k} = \begin{bmatrix} J_k(1) & U_k \\ 0 & J_k(1)^{-\top} \end{bmatrix},$$

where the $(1,2)$-block U_k is given by $U_k = (0_{k-1,k-1} \oplus [1])J_k(1)^{-\top}$. The matrix F_{2k} is similar to $J_{2k}(1)$. We have the following canonical form for symplectic matrices under symplectic similarity [7, Lemma 5].

Theorem 2.14. (de la Cruz, 2015) *Let $A \in \mathrm{Sp}(2n, \mathbb{C})$. Then A is symplectically similar to an expanding sum of symplectic matrices of the form*

1. $J_k(\lambda) \oplus J_k(\lambda)^{-\top}$, *where* $\lambda \in \mathbb{C}\backslash\{-1, 0, 1\}$ *and* $k \in \mathbb{N}$,
2. $J_{2k-1}(\lambda) \oplus J_{2k-1}(\lambda)^{-\top}$, *where* $\lambda \in \{-1, 1\}$ *and* $k \in \mathbb{N}$, *and*
3. $\pm F_{2k}$, *where* $k \in \mathbb{N}$.

Using Theorems 2.14 and 2.2, we construct a canonical form for symplectic matrices under * congruence. Let $A \in \mathrm{Sp}(2n, \mathbb{C})$. Then $A^{-*}A \in \mathrm{Sp}(2n, \mathbb{C})$. In this case,

$$A^{-*}A = (J^{-1}A^\top J)^* A = J^{-1}\overline{A}JA = -\overline{JA}(JA).$$

The Jordan Canonical Form of the matrix $\overline{X}X$ has a special structure. By [24, Corollary 4.6.16], $\overline{X}X$ is similar to the square of a real matrix. As a consequence, if $\lambda < 0$ is an eigenvalue of $\overline{X}X$, then the number of blocks of $J_k(\lambda)$ is even for all $k \in \mathbb{N}$. This implies that the number of blocks $J_k(\lambda)$, where $\lambda > 0$, in the Jordan Canonical Form of $A^{-*}A$ is even. Using this fact, the symplectic Jordan structure, and the Jordan structure of a general *cosquare, we obtain that $A^{-*}A$ is similar to an expanding sum of symplectic matrices of the form

1. $J_k(\mu) \oplus J_k(\mu)^{-\top}$, where $\mu < -1$,
2. $J_k(\mu) \oplus J_k(\mu^{-1}) \oplus J_k(\mu)^{-\top} \oplus J_k(\mu^{-1})^{-\top}$, where $\mu > 1$,
3. $J_k(\mu) \oplus J_k(\overline{\mu^{-1}}) \oplus J_k(\mu)^{-\top} \oplus J_k(\overline{\mu^{-1}})^{-\top}$, where $\mu \in \mathbb{C}\backslash\mathbb{R}$ and $|\mu| > 1$,
4. $J_k(\lambda) \oplus J_k(\lambda)^{-\top}$, where $\lambda \in \mathbb{C}\backslash\{-1, 1\}$ and $|\lambda| = 1$,
5. $J_{2k-1}(\lambda) \oplus J_{2k-1}(\lambda)^{-\top}$, where $\lambda \in \{-1, 1\}$, and
6. $\pm F_{2k}$.

For each block type listed above, we construct a corresponding block for symplectic *congruence.

1. We define the first canonical block $E_{2k}(\mu)$ to be the symplectic block antidiagonal matrix

$$E_{2k}(\mu) = \begin{bmatrix} 0 & -J_k(\sqrt{-\mu})^{-\top} \\ J_k(\sqrt{-\mu}) & 0 \end{bmatrix}.$$

   Observe that

$$\begin{aligned} & E_{2k}(\mu)^{-*}E_{2k}(\mu) \\ = & \begin{bmatrix} 0 & J_k(\sqrt{-\mu})^{-1} \\ -J_k(\sqrt{-\mu})^{\top} & 0 \end{bmatrix}^{\top} \begin{bmatrix} 0 & -J_k(\sqrt{-\mu})^{-\top} \\ J_k(\sqrt{-\mu}) & 0 \end{bmatrix} \\ = & \begin{bmatrix} 0 & -J_k(\sqrt{-\mu}) \\ J_k(\sqrt{-\mu})^{-\top} & 0 \end{bmatrix} \begin{bmatrix} 0 & -J_k(\sqrt{-\mu})^{-\top} \\ J_k(\sqrt{-\mu}) & 0 \end{bmatrix} \\ = & \begin{bmatrix} -J_k(\sqrt{-\mu})^2 & 0 \\ 0 & (-J_k(\sqrt{-\mu})^2)^{-\top} \end{bmatrix}. \end{aligned}$$

   The matrix $-J_k(\sqrt{-\mu})^2$ is similar to $J_k(\mu)$ so $E_{2k}(\mu)^{-*}E_{2k}(\mu)$ is similar to $J_k(\mu) \oplus J_k(\mu)^{-*}$. By Lemma 2.1, there exist matrices D_- and D_+ such that $E_{2k}(\mu)$ is *congruent to $-D_- \oplus D_+$ and $H_{2k}(\mu)$ is *congruent to $D_- \oplus D_+$. Since $H_{2k}(\mu)$ is already a canonical block, one of D_- or D_+ must be absent. Thus, $E_{2k}(\mu)$ is *congruent to $\pm H_{2k}(\mu)$. But $-H_{2k}(\mu)$ is *congruent to $H_{2k}(\mu)$ via the matrix

$$\begin{bmatrix} I_k & 0 \\ 0 & -I_k \end{bmatrix}.$$

   Thus, $E_{2k}(\mu)$ is *congruent to $H_{2k}(\mu)$.

2. Since we have a quadruple of blocks, the corresponding *congruence canonical block is taken to be $H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$.

3. Same as in (2).

4. We choose the *congruence canonical block to be of the form $\mu D_k \oplus (\mu D_k)^{-\top}$, where $|\mu| = 1$.

5. Same as in (4).

6. The last canonical block G_{2k} is a symplectic matrix A such that $A^{-*}A$ that is similar to $-F_{2k}$. We take

$$G_{2k} = J_{2k}^{-1} F_{2k} = \begin{bmatrix} 0 & -J_k(1)^{-\top} \\ J_k(1) & U_k \end{bmatrix}.$$

Then

$$G_{2k}^{-*} G_{2k} = G_{2k}^{-\top} G_{2k} = J_{2k}^{T} F_{2k}^{-\top} J_{2k}^{-1} F_{2k} = -J_{2k}^{-1} F_{2k}^{-\top} J_{2k} F_{2k} = -F_{2k}^{2},$$

which is similar to $-F_{2k}$. Let $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$. Note that $\lambda G_{2k} \in \mathrm{Sp}(2k, \mathbb{C})$ if and only if $\lambda = \pm 1$.

Theorem 2.15. *Let $A \in \mathrm{Sp}(2n, \mathbb{C})$. Then A is symplectically *congruent to an expanding sum of symplectic matrices of the form*

1. *$E_{2k}(\mu)$, where $\mu < 0$, $|\mu| > 1$, $k \in \mathbb{N}$,*

2. *$H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$, where $|\mu| > 1$, $k \in \mathbb{N}$,*

3. *$\lambda D_k \oplus (\lambda D_k)^{-\top}$, where $|\lambda| = 1$, $k \in \mathbb{N}$, and*

4. *$\pm G_{2k}$, where $k \in \mathbb{N}$.*

We may replace the pair $E_{2k}(\mu) \boxplus E_{2k}(\mu)$ by $H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$ and the pair $G_{2k} \boxplus G_{2k}$ by $D_{2k} \oplus D_{2k}^{-\top}$.

Chapter 3

Proof of Theorem 1.1.

It suffices to look at each canonical block and show that each is a product of five positive definite symplectic matrices.

3.0.1 Type I block $E_{2k}(\mu)$

Let $\mu < -1$. Let $D, E \in M_k(\mathbb{C})$ be diagonal matrices. We can write

$$E_{2k}(\mu) = \begin{bmatrix} J_k(\sqrt{-\mu})^{-\top} D & 0 \\ J_k(\sqrt{-\mu})E & J_k(\sqrt{-\mu})D^{-1} \end{bmatrix} \begin{bmatrix} 0 & -D^{-1} \\ D & E \end{bmatrix}.$$

Since D, E are diagonal matrices, the second factor is symplectic. Since the product is symplectic, the first factor must be symplectic as well. We choose D to have distinct positive diagonal entries such that none are pairwise reciprocal. Then the first factor is symplectically similar to an expanding sum of 2×2 symplectic matrices with distinct positive eigenvalues. Hence, the first factor is a product of two positive definite symplectic matrices. The second factor has characteristic polynomial

$$f(\lambda) = \det(\lambda^2 I_k - \lambda E + I_k) = \prod_{j=1}^{n} (\lambda^2 - e_{jj}\lambda + 1).$$

We choose E to have distinct diagonal entries such that $e_{jj} > 4$. This means that the quadratic equation $\lambda^2 - e_{jj}\lambda + 1$ has positive roots. With our choice of E, the eigenvalues of the second factor are distinct and positive. Therefore, the second factor is also a product of two positive definite symplectic matrices.

Proposition 3.1. *Let $A = E_{2k}(\mu)$ for some $\mu < -1$. Then A is a product of four positive definite symplectic matrices.*

3.0.2 Type II block $H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$

We consider the canonical block $H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$, where $|\mu| > 1$, and $k \in \mathbb{N}$. Let $1 < a_1 < \ldots < a_n$ and consider $D = \mathrm{diag}(a_1, \ldots, a_n)$. We can write $H_{2n}(\mu)$ as $H_{2n}(\mu) = H_1 H_2$, where

$$H_1 = I \oplus D, \quad H_2 = \begin{bmatrix} 0 & I \\ H_3 & 0 \end{bmatrix}.$$

The matrix H_3 is upper triangular of the form

$$H_3 = \begin{bmatrix} a_1^{-1}\mu & a_1^{-1} & & \\ & a_2^{-1}\mu & \ddots & \\ & & \ddots & a_{n-1}^{-1} \\ & & & a_n^{-1}\mu \end{bmatrix}.$$

The matrix H_2 has $2n$ distinct eigenvalues so H_2 is diagonalizable. We choose the a_j's so that $-1 \notin \sigma(H_2)$. It follows that $H_2 \oplus H_2^{-\top}$ is symplectically similar to an expanding sum of 2×2 nonscalar symplectic matrices. Hence, by Corollary 2.2, $H_2 \oplus H_2^{-\top}$ is a product of four positive definite symplectic matrices. The matrix $H_1 \oplus H_1^{-\top}$ is positive definite symplectic.

Proposition 3.2. *Let $A = H_{2k}(\mu) \oplus H_{2k}(\mu)^{-\top}$ for some $|\mu| > 1$ and $k \in \mathbb{N}$. Then A is a product of five positive definite symplectic matrices.*

3.0.3 Type III block $\lambda D_k \oplus (\lambda D_k)^{-\top}$

In this section, we consider the canonical block $\lambda D_k \oplus \lambda^{-1} D_k^{-\top}$, where $|\lambda| = 1$ and $k \in \mathbb{N}$. We are now ready to show that $\lambda D_k \oplus (\lambda D_k)^{-\top}$ is a product of five positive definite symplectic matrices. By the previous proposition, D_k has distinct eigenvalues. Choose $p > 0$ such that

$-1 \notin \sigma(p^{-1}\lambda D_k)$. Let $P = pI_k \oplus p^{-1}I_k$, which is positive definite symplectic. Then

$$\lambda D_k \oplus \lambda^{-1} D_k^{-\top} = P(p^{-1}\lambda D_k \oplus (p^{-1}\lambda D_k)^{-\top}).$$

The second factor is symplectically similar to an expanding sum of 2×2 symplectic matrices, none of which is $-I$. It follows that the second factor is a product of four positive definite symplectic matrices.

Proposition 3.3. *Let $A = \lambda D_k \oplus (\lambda D_k)^{-\top}$ for some $|\lambda| = 1$ and $k \in \mathbb{N}$. Then A is a product of five positive definite symplectic matrices.*

3.0.4 Type IV block $\pm G_{2k}$

Let $k \in \mathbb{N}$ and consider the canonical matrix G_{2k}. Let

$$p_k(\lambda) = \det(\lambda I - G_{2k})$$

be its characteristic polynomial. Observe that $p_1(\lambda) = \lambda^2 - \lambda + 1$ and for $k > 1$,

$$p_k(\lambda) = \det \begin{bmatrix} \lambda I_k & J_k(1)^{-\top} \\ -J_k(1) & \lambda I_k - U_k \end{bmatrix} = \det(\lambda^2 I_k - \lambda U_k + J_k(1)J_k(1)^{-\top}).$$

A direct computation reveals that $J_k(1)J_k(1)^{-\top}$ is a companion matrix:

$$J_k(1)J_k(1)^{-\top} = \begin{bmatrix} 0 & 1 & & \\ \vdots & & 1 & \\ 0 & & & 1 \\ (-1)^{k+1} & \cdots & -1 & 1 \end{bmatrix}.$$

Now,

$$\begin{aligned}
p_k(\lambda) &= \det \begin{bmatrix} \lambda^2 & e_1^\top \\ (-1)^k(\lambda-1)e_{k-1} & \lambda^2 I_{k-1} - \lambda U_{k-1} + J_{k-1}(1)J_{k-1}(1)^{-\top} \end{bmatrix} \\
&= \lambda^2 p_{k-1}(\lambda) + (-1)^{2k+1}(\lambda-1) \\
&= \lambda^2 p_{k-1}(\lambda) - \lambda + 1.
\end{aligned}$$

Computing the first few k values, we see the pattern more clearly:

$$\begin{aligned}
p_1(\lambda) &= \lambda^2 - \lambda + 1 \\
p_2(\lambda) &= \lambda^2(\lambda^2 - \lambda + 1) - \lambda + 1 = \lambda^4 - \lambda^3 + \lambda^2 - \lambda + 1 \\
p_3(\lambda) &= \lambda^2 p_2(\lambda) - \lambda + 1 = \lambda^6 - \lambda^5 + \lambda^4 - \lambda^3 + \lambda^2 - \lambda + 1.
\end{aligned}$$

In general, $p_k(\lambda) = \lambda^{2k} - \lambda^{2k-1} + \cdots + \lambda^2 - \lambda + 1$. The roots of p_k are the $2k$ distinct nonreal $(2k+1)th$ roots of -1.

Proposition 3.4. *The matrix G_{2k} has distinct nonreal eigenvalues for $k \in \mathbb{N}$.*

Note that $\pm 1 \notin \sigma(G_{2n})$. Thus, $\pm G_{2k}$ is similar to an expanding sum of 2×2 nonscalar symplectic matrices. Thus, $\pm G_{2k}$ is a product of four positive definite symplectic matrices.

Proposition 3.5. *Let $A = \pm G_{2k}$ for some $k \in \mathbb{N}$. Then A is a product of four positive definite symplectic matrices.*

Chapter 4

Future research

The complex symplectic group $\mathrm{Sp}(2n, \mathbb{C})$ is one out of a myriad of matrix groups with positive determinant. The closely related real symplectic group $\mathrm{Sp}(2n, \mathbb{R})$ inherits many of the linear algebraic properties that have been useful in solving the complex case. However, due to $\mathbb{R}$ not being algebraically closed, canonical forms over $\mathbb{R}$ tend to have more complicated block structures. Gutt [20] gave a canonical form for $\mathrm{Sp}(2n, \mathbb{R})$, but this seems to be impractical for the methods used in this work. We describe the canonical blocks below and explain some of the difficulties in exploiting these structures. Let $\mathbb{T} = \{e^{i\theta} : \theta \in \mathbb{R}\}$, the unit circle in $\mathbb{C}$. We define the following canonical matrices:

1. For $k \in \mathbb{N}$, $J_k(\lambda)$ denotes the $k \times k$ upper triangular Jordan matrix corresponding to $\lambda \in \mathbb{C}$,

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

2. For $\theta \in \mathbb{R}$ and $r \geq 0$, we denote the rotation matrix by

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Let $\lambda = re^{i\theta} \in \mathbb{C}$. We denote by $R_k(\lambda)$ the $2k \times 2k$ block upper triangular matrix

$$R_k(\lambda) = \begin{bmatrix} rR(\theta) & I_2 & & & \\ & rR(\theta) & I_2 & & \\ & & \ddots & \ddots & \\ & & & rR(\theta) & I_2 \\ & & & & rR(\theta) \end{bmatrix}.$$

3. Let $s \in \{-1, 1\}$ and $\lambda \in \mathbb{C}$. We define $S_k(\lambda, s)$ to be the $2k \times 2k$ block upper triangular matrix

$$S_k(\lambda, s) = \begin{bmatrix} J_k(\lambda)^{-1} & J_k(\lambda)^{-1}(0 \oplus [s]) \\ 0 & J_k(\lambda)^\top \end{bmatrix}.$$

4. Let $s \in \{-1, 1\}$ and $\lambda \in \mathbb{C}$. We define $U_k(\lambda, s)$ to be the $4k \times 4k$ block upper triangular matrix

$$U_k(\lambda, s) = \begin{bmatrix} R_k(\overline{\lambda})^{-1} & R_k(\overline{\lambda})^{-1}(0 \oplus sI_2) \\ 0 & R_k(\overline{\lambda})^\top \end{bmatrix}.$$

5. Let $s \in \{-1, 1\}$ and $\lambda = e^{i\theta} \in \mathbb{C}$. Denote by e_j the jth standard basis vector of appropriate size. We define

$$V_0(\lambda, s) = \begin{bmatrix} \cos\theta & s\sin\theta \\ -s\sin\theta & \cos\theta \end{bmatrix}$$

and for $k \in \mathbb{N}$, $V_k(\lambda, s)$ is the $(4k + 2) \times (4k + 2)$ block upper triangular matrix

$$V_k(\lambda, s) = \begin{bmatrix} R_k(\overline{\lambda})^{-1} & sy & R_k(\overline{\lambda})^{-1}(0 \oplus \frac{s}{2}J_2^{-1}) & x \\ 0 & \cos\theta & e_{k-1}^\top & s\sin\theta \\ 0 & 0 & R_k(\overline{\lambda})^\top & 0 \\ 0 & -s\sin\theta & -se_k^\top & \cos\theta \end{bmatrix},$$

where $x = R_k(\overline{\lambda})^{-1}(0 \oplus R(\theta))e_{k-1}$ and $y = R_k(\overline{\lambda})^{-1}(0 \oplus R(\theta))e_k$.

The following theorem gives a canonical form for real symplectic matrices under real symplectic similarity.

Theorem 4.1 (Real Symplectic Canonical Form, Gutt 2014). *Every matrix in* $\mathrm{Sp}(2n, \mathbb{R})$ *is real symplectically similar to an expanding sum of the following canonical blocks.*

1. $J_k(\lambda)^{-1} \oplus J_k(\lambda)^\top$, *where* $\lambda \in \mathbb{R}\backslash\{-1, 0, 1\}$, $k \in \mathbb{N}$,
2. $J_k(\lambda)^{-1} \oplus J_k(\lambda)^\top$, *where* $\lambda \in \{-1, 1\}$, k *is odd,*
3. $R_k(\lambda)^{-1} \oplus R_k(\lambda)^\top$, *where* $\lambda \in \mathbb{C}\backslash(\mathbb{T} \cup \mathbb{R})$, $k \in \mathbb{N}$,
4. $S_k(\lambda, s)$, *where* $\lambda \in \{-1, 1\}$, $s \in \{-1, 1\}$, $k \in \mathbb{N}$,
5. $U_k(\lambda, s)$, *where* $\lambda \in \mathbb{T}\backslash\{-1, 1\}$, $s \in \{-1, 1\}$, $k \in \mathbb{N}$,
6. $V_k(\lambda, s)$, *where* $\lambda \in \mathbb{T}\backslash\{-1, 1\}$, $s \in \{-1, 1\}$, $k \in \mathbb{N} \cup \{0\}$.

The expanding sum is unique up to permutation of canonical blocks and the choice of representative eigenvalues coming from the set $[\lambda] = \{\lambda, \lambda^{-1}, \overline{\lambda}, \overline{\lambda^{-1}}\}$.

If we follow the strategy we used to solve the complex symplectic case. We use Theorem 4.1 to construct a real symplectic *congruence canonical form. This requires a considerable amount of time beyond the timeline set for this dissertation. Another difficulty that arises from the real cases is the existence of real square roots of real matrices, and more generally, the existence of real polynomial interpolations for matrix functions. Existence of real matrices are tied to having certain Jordan block structures.

The 2×2 orthogonal group $\mathrm{O}(2, \mathbb{C})$ is an example of a group that cannot be expressed as a product of five (or more) positive definite orthogonal matrices. By [27, Theorem 6], a complex orthogonal $Q \in \mathrm{O}(2, \mathbb{C})$ can be written as $Q = e^S$ for some skew-symmetric $S \in M_2(\mathbb{C})$ if and only if

$$Q \neq \pm \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Any skew-symmetric $S \in M_2(\mathbb{C})$ can be written as $S = zJ_2$ for some $z \in \mathbb{C}$. Consequently, $Q = e^{zJ_2}$ has positive eigenvalues if and only if z is purely imaginary. Moreover, if z is purely imaginary, Q is positive definite since $Q^* = Q$. Every 2×2 skew-symmetric matrix is a scalar multiple of J_2, thus, any pair of skew-symmetric matrices $S_1, S_2 \in M_2(\mathbb{C})$ commute. It follows that $e^{S_1}e^{S_2} = e^{S_1+S_2}$ is satisfied whenever we are given 2×2 skew-symmetric matrices S_1 and S_2. We conclude that the set of all 2×2 positive definite complex orthogonal matrices is closed under the usual matrix multiplication. The case $n = 2$ might be an exception and it is still worth looking at the orthogonal groups $\mathrm{O}(n, \mathbb{C})$ with $n > 2$ and determine what are necessary and sufficient conditions under which the problem is solved. We note that the complex orthogonal group $\mathrm{O}(n, \mathbb{C})$ suffers from the same problem as $\mathrm{Sp}(2n, \mathbb{R})$: there are no clear canonical blocks with good block structure to work with.

A future direction based on this work is to consider $\mathrm{Sp}(2n, \mathbb{R})$ or the complex orthogonal group $\mathrm{O}(n, \mathbb{C})$ but with a different, more general approach. Both $\mathrm{Sp}(2n, \mathbb{C})$ and $\mathrm{O}(n, \mathbb{C})$ can be described in terms of *antihomomorphic, spectrum-preserving linear maps* ϕ defined by

$$\phi(A) = S^{-1}A^{\top}S,$$

where $S = J$ in the case of $\mathrm{Sp}(2n, \mathbb{C})$ and $S = I$ in the case of $\mathrm{O}(n, \mathbb{C})$. The author has done an extensive work in this subject in the past [8, 14, 15, 16, 17]; using tools and ideas from this area proves to be promising. As an example, the proof of Theorem 2.12 remains valid if we replace $\mathrm{Sp}(2n, \mathbb{C})$ by $\mathrm{O}(n, \mathbb{C})$ and all instances of A^J are replaced by $A^{\top}$ or more generally, $\phi(A)$ for some nonsingular S with good properties.

References

[1] C.S. Ballantine, Products of positive definite matrices. I, Pacific J. Math. **23** (1967) 427-433.

[2] C.S. Ballantine, Products of positive definite matrices. II, Pacific J. Math. **24** (1968) 7-17.

[3] C.S. Ballantine, Products of positive definite matrices. III, J. Algebra **10** (1968) 174-182.

[4] C.S. Ballantine, Products of positive definite matrices. IV, Linear Algebra Appl. **3** (1970) 79-114.

[5] J.D. Botha, A unification of some matrix factorization results, Linear Algebra Appl. **431** (2009) 1719-1725.

[6] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Algebraic Geometry and Commutative Algebra*, 2nd Edition, New York, Springer-Verlag, 1996.

[7] R.J. de la Cruz, Each symplectic matrix is a product of four symplectic involutions, Linear Algebra Appl. **466** (2015) 382-400.

[8] R.J. de la Cruz and D.Q. Granario, The ϕ_S polar decomposition when the cosquare of S is nonderogatory. Electron. J. Linear Algebra **31** (2016) 754-764.

[9] J. Cui, C.-K. Li, and N.-S. Sze, Products of positive semi-definite matrices, Linear Algebra Appl. 528 (2017) 17-24.

[10] K. Dilcher and K.B. Stolarsky, Resultants and discriminants of Chebyshev and related polynomials, Trans. Amer. Math. Soc. **357** (2004) 965-981.

[11] L. Elsner and Kh.D. Ikramov, Normal matrices: an update, Linear Algebra Appl. **285** (1998) 291-303.

[12] H. Faßbender and Kh.D. Ikramov, Several observations on symplectic, Hamiltonian, and skew-Hamiltonian matrices, Linear Algebra Appl. **400** (2005) 15-29.

[13] H. Faßbender and Kh.D. Ikramov, A note on an unusual type of polar decomposition, Linear Algebra Appl. **429** (2008) 42-49.

[14] D.Q. Granario, D.I. Merino, and A.T. Paras, The ϕ_S polar decomposition. Linear Algebra Appl. **438** (2013) 609-620.

[15] D.Q. Granario, D.I. Merino, and A.T. Paras, The ϕ_S polar decomposition when the cosquare of S is normal. Linear Algebra Appl. **467** (2015) 75-85.

[16] D.Q. Granario, D.I. Merino, and A.T. Paras, The ψ_S-polar decomposition when the cosquare of S is normal. Linear Algebra Appl. **495** (2016) 51-66.

[17] D.Q. Granario, D.I. Merino, and A.T. Paras, The sum of two ϕ_S orthogonal matrices when $S^{-\top}S$ is normal and $-1 \notin \sigma(S^{-\top}S)$. Linear Algebra Appl. **495** (2016) 67-89.

[18] D. Q. Granario and T.-Y. Tam, Products of positive definite symplectic matrices, submitted for publication.

[19] R. Grone, C.R. Johnson, E.M. Sa, and H. Wolkowicz, Normal matrices, Linear Algebra Appl. **87** (1987) 213-225.

[20] J. Gutt, Normal forms for symplectic matrices, Port. Math. **71** (2014) 109-139.

[21] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, American Mathematical Society, Providence, RI, 2001. Corrected reprint of the 1978 original.

[22] N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, PA, 2008.

[23] N.J. Higham, D.S. Mackey, N. Mackey, and F. Tisseur, Functions preserving matrix groups and iterations for the matrix square root, SIAM J. Matrix Anal. Appl. **26** (2005) 849-877.

[24] R.A. Horn and C.R. Johnson, *Matrix Analysis*, 2nd Edition, Cambridge University Press, New York, NY, 2013.

[25] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, NY, 1991.

[26] R.A. Horn and D.I. Merino, Contragredient equivalence: a canonical form and some applications, Linear Algebra Appl. **214** (1995) 43-92.

[27] R.A. Horn and D.I. Merino, The Jordan canonical forms of complex orthogonal and skew-symmetric matrices, Linear Algebra Appl. **302-303** (1999) 411-421.

[28] R.A. Horn and V.V. Sergeichuk, Canonical forms for complex matrix congruence and conjunctivity, Linear Algebra Appl. **416** (2006) 1010-1032.

[29] B. Kostant, On convexity, the Weyl group and the Iwasawa decomposition. Ann. Sci. École Norm. Sup. **(4) 6** (1974) 413-455.

[30] D. Kulkarni, D. Schmidt, and S.-K. Tsui, Eigenvalues of tridiagonal pseudo-Toeplitz matrices, Linear Algebra Appl. **297** (1999) 63-80.

[31] D.S. Mackey, N. Mackey, and F. Tisseur, Structured factorizations in scalar product spaces, SIAM J. Matrix Anal. Appl. **27** (3) (2006) 821-850.

[32] G.J. Murphy, Products of positive operators, Proc. Amer. Math. Soc. **125** (1997) 3675-3677.

[33] N.C. Phillips, Every invertible Hilbert space operator is a product of seven positive operators, Canad. Math. Bull. **38** (1995) 230-236.

[34] H. Radjavi, Products of Hermitian matrices and symmetries, Proc. Amer. Math Soc. **21** (1969) 369-372; Errata, Proc. Amer. Math. Soc. **26** (1970) 701.

[35] D. Serre, *Matrices: Theory and Applications,* 2nd edition, Springer-Verlag, New York, 2010.

[36] A.R. Sourour, A factorization theorem for matrices, Linear and Multilinear Algebra **19** (1986) 141-147.

[37] O. Taussky, Problem 4846, Proc. Amer. Math Soc. **66** (1959) 427.

[38] P.Y. Wu, Products of normal operators, Canad. J. Math. **40** (1988) 1322-1330.

[39] P.Y. Wu, Products of positive semidefinite matrices, Linear Algebra Appl. **111** (1988) 53-61.

[40] P.Y. Wu, The operator factorization problems, Linear Algebra Appl. **117** (1989) 35-63.

[41] D. Zwillinger, *CRC Standard Mathematical Tables and Formulae*, 31st Edition, CRC Press, Boca Raton, FL, 2003.