Publications

2-15-2019

Action of Complex Symplectic Matrices on the Siegel Upper Half Space

Keshav R. Acharya
Embry-Riddle Aeronautical University, acharyak@erau.edu

Matt McBride
Mississippi State University

ACTION OF COMPLEX SYMPLECTIC MATRICES ON THE SIEGEL UPPER HALF SPACE

KESHAV RAJ ACHARYA AND MATT MCBRIDE

ABSTRACT. The Siegel upper half space, $\mathcal{S}_n$, the space of complex symmetric matrices, Z with positive definite imaginary part, is the generalization of the complex upper half plane in higher dimensions. In this paper, we study a generalization of linear fractional transformations, Φ_S, where S is a complex symplectic matrix, on the Siegel upper half space. We partially classify the complex symplectic matrices for which $\Phi_S(Z)$ is well defined. We also consider $\mathcal{S}_n$ and $\overline{\mathcal{S}}_n$ as metric spaces and discuss distance properties of the map Φ_S from $\mathcal{S}_n$ to $\mathcal{S}_n$ and $\overline{\mathcal{S}}_n$ respectively.

Keywords: Siegel upper half space, Symplectic matrix, Finsler metric

AMS (MOS) Subject Classification. 15A60.

1. INTRODUCTION

The Siegel upper half-space which we denote by $\mathcal{S}_n$ is the set of all $n \times n$ symmetric matrices Z of the form: $Z = X + iY$ where X and Y are symmetric matrices and Y is positive definite. When $n = 1$, then

$$\mathcal{S}_1 = \mathbb{C}^+ = \{z \in \mathbb{C} : z = x + iy, \ y > 0\} \ .$$

Therefore the space $\mathcal{S}_n$ is a generalization of the complex upper half plane. For more on details on the Siegel space see [8]. Similarly, we consider the symplectic group $SP_{2n}(\mathbb{C})$ as a generalization of $SL_2(\mathbb{R})$, the group of all $2n \times 2n$ complex matrices S satisfying $S^t J S = J$ with

$$J = \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix}$$

where I_n is the $n \times n$ identity matrix and O is the $n \times n$ zero matrix.

An action of $SL_2(\mathbb{R})$ on $\mathbb{C}^+$ is given by fractional transformations,

$$T(z) = \frac{az + b}{cz + d} \tag{1.1}$$

where $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$, which has been an useful tool in the theory of functions of complex variable.

To relate the symplectic group to the Siegel upper half space, we write a symplectic matrix $S \in SP_{2n}(\mathbb{C})$ in the form of block matrices

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A, B, C, D are $n \times n$ complex matrices. It can be easily seen by using the equation $S^t J S = J$ that S is symplectic if and only if $A^t C$ and $B^t D$ are symmetric and $A^t D - C^t B = I_n$.

The generalization of the fractional transformation given by equation (1.1) to fractional transformations on $\mathcal{S}_n$, is given by a map

$$\Phi_S(Z) = S(Z) = (AZ + B)(CZ + D)^{-1}, \ Z \in \mathcal{S}_n. \qquad (1.2)$$

The map $\Phi_S(Z)$ is well defined if and only if $CZ + D$ is invertible.

In [3, 8], it is shown that if $S \in SP_{2n}(\mathbb{R})$ and $Z \in \mathcal{S}_n$ then $\Phi_S(Z)$ is well defined and it maps the Siegel upper half space to itself. Thus we can consider the map

$$\Phi : SP_{2n}(\mathbb{R}) \times \mathcal{S}_n \to \mathcal{S}_n \text{ via } (S, Z) \mapsto \Phi_S(Z). \qquad (1.3)$$

Since $SP_{2n}(\mathbb{R})$ is a group, Φ has been considered as a group action of $SP_{2n}(\mathbb{R})$ on $\mathcal{S}_n$.

This action was first studied by Carl Ludwig Siegel in 1943 and published his work in the book "Symplectic Geometry" in which not only the analytical and geometrical aspects of the action are considered but also some applications to number theory. In 1999, Freitas gave more insight into this theory from the point of view of topological spaces, see [3] for details. In particular he compactified the space and extended the action continuously to a compact domain. There are several studies on the Siegel upper half space motivated from other areas, for example [6, 7].

A natural question to ask is: can the map Φ be extended to a map $\tilde{\Phi}$ given by

$$\tilde{\Phi} : SP_{2n}(\mathbb{C}) \times \mathcal{S}_n \to \mathcal{S}_n \text{ via } (S, Z) \mapsto \Phi_S(Z)?$$

The following simple example illustrates that $\tilde{\Phi}$ can not be defined for all $S \in SP_{2n}(\mathbb{C})$.

Example 1.1. *Let* $S = \begin{pmatrix} -iI_n & 0 \\ -I_n & iI_n \end{pmatrix}$ *then* $S \in SP_{2n}(\mathbb{C})$ *but* $\Phi_S(iI_n)$ *is not defined as* $(-iI_n + iI_n)$ *is not invertible .*

In addition, even if the map $\Phi_S(Z)$ is well defined, it is possible that for some $S \in SP_{2n}(\mathbb{C})$, $\Phi_S(Z) \notin \mathcal{S}_n$ as in the example below.

Example 1.2. *Let* $S = \begin{pmatrix} I_n & iI_n \\ iI_n & O \end{pmatrix}$. *Notice that* $S \in SP_{2n}(\mathbb{C})$ *and* $iI_n \in \mathcal{S}_n$. *Therefore*

$$\Phi_S(iI_n) = (iI_n + iI_n)(iI_n iI_n)^{-1} = -2iI_n .$$

Thus $\Phi_S(iI_n) \notin \mathcal{S}_n$.

These two examples show that such an extension can not be arbitrarily made and one needs to focus on a proper domain for $\tilde{\Phi}$.

Let $A(2n, \mathbb{C})$ be the space of all matrices S that satisfy $S^t JS = -J$, that is S is antisymplectic. Define the following space:

$$\overline{\mathcal{S}}_n = \{Z \in \mathbb{C}^{d \times d} : Z = X + iY, \ X^t = X, Y^t = Y, \ Y < 0\} .$$

Following the methods of [3], it is easy to see that if $S \in A(2n, \mathbb{R})$, Φ_S maps the Siegel half space to the space $\overline{\mathcal{S}}_n$. We will show this in Section 2. Since the product of antisymplectic matrices is a symplectic matrix, the set $A(2n, \mathbb{C})$ is not a group. Also, if we consider S to be antisymplectic complex matrix, Φ_S does not map $\mathcal{S}_n$ to $\overline{\mathcal{S}}_n$.

Example 1.3. *Let* $S = \begin{pmatrix} iI_n & O \\ O & iI_n \end{pmatrix}$ *then* $S \in A(2n, \mathbb{C})$ *and* $iI_n \in \mathcal{S}_n$. *However* $\Phi_S(iI_n) \notin \overline{\mathcal{S}}_n$ *since*

$$\Phi_S(iI_n) = (iI_n iI_n)(iI_n)^{-1} = iI_n .$$

One of the main goals in this paper is to partially classify all complex symplectic matrices for which the map $\Phi_S : \mathcal{S}_n \to \mathcal{S}_n$ is well defined on $\mathcal{S}_n$.

We also want to consider a metric on $\mathcal{S}_n$ as a generalization of the hyperbolic metric

$$ds = \frac{\sqrt{dx^2 + dy^2}}{y} \tag{1.4}$$

on the complex hyperbolic plane $\mathbb{C}^+$. For this we consider the Finsler metric as in [2] as follows.

Let $Z = X + iY \in \mathcal{S}_n$ and let the complex symmetric matrix W be an element of the tangent space at Z. Consider the Finsler norm F_Z defined by

$$F_Z(W) = \|Y^{-\frac{1}{2}} W Y^{-\frac{1}{2}}\|$$

where $\|\cdot\|$ is the usual matrix (operator) norm. This Finsler norm defines a metric on $\mathcal{S}_n$ as

$$d_\infty(Z_1, Z_2) = \inf_{Z(t)} \int_0^1 F_{Z(t)}(\dot{Z}(t)) \ dt$$

where the infimum is taken over all differentiable paths $Z(t)$ joining Z_1 to Z_2.

In the case when $n = 1$, the Siegel upper half space $\mathcal{S}_1 = \mathbb{C}^+$ and the metric d_∞ is same as the the hyperbolic metric defined in equation (1.4) on $\mathbb{C}^+$. The length of a curve in the Euclidean plane is measured by $\sqrt{dx^2 + dy^2}$. Similarly, the length of a curve $\gamma(t) = x(t) + iy(t)$, $t \in [0, 1]$ on the hyperbolic plane $\mathbb{C}^+$ is defined by

$$h(\gamma) = \int_0^1 \frac{1}{y(t)} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \, dt \ .$$

The hyperbolic distance between two points $z, w \in C^+$ is defined by

$$\rho(z, w) = \inf h(\gamma)$$

where the infimum is taken over all piecewise differentiable curves joining z and w. So for $n = 1$, we have $d_\infty(z, w) = \rho(z, w)$. Thus the Finsler metric d_∞ on $\mathcal{S}_n$ is a generalization of the hyperbolic metric ρ on $\mathbb{C}^+$.

We also want to define a metric on $\mathcal{S}_n$. First note that the complex conjugate of a given matrix $Z = X + iY \in \overline{\mathcal{S}}_n$ is a matrix $\overline{Z} = X - iY \in \mathcal{S}_n$. A metric d_- on $\overline{\mathcal{S}}_n$ is defined by

$$d_-(Z_1, Z_2) = d_\infty(\overline{Z}_1, \overline{Z}_2) \ .$$

Thus $(\mathcal{S}_n, d_\infty)$ and $(\overline{\mathcal{S}}_n, d_-)$ are metric spaces.

2. Preliminaries and Results

As mentioned above, it is shown in [3] that for each real symplectic matrix S, the map Φ_S maps the Siegel upper half space to itself. However, it is not true when S is real antisymplectic matrix. In fact, it maps $\mathcal{S}_n$ onto $\overline{\mathcal{S}}_n$. We have the following proposition.

Proposition 2.1. *For $S \in A(2n, \mathbb{R})$ the map Φ_S maps $\mathcal{S}_n$ onto $\overline{\mathcal{S}}_n$.*

Proof. We show that $CZ + D$ is invertible for any $Z \in \mathcal{S}_n$ and $\Phi_S(Z) = (AZ + B)(CZ + D)^{-1} \in \overline{\mathcal{S}}_n$.

Set $E = AZ + B$ and $F = CZ + D$. First we show that F is invertible. We can write $\Phi_S(Z)$ as

$$S \begin{pmatrix} Z \\ I_n \end{pmatrix} = \begin{pmatrix} E \\ F \end{pmatrix} \ .$$

From which we have,

$$\frac{1}{2i}(E^* \; F^*)J\begin{pmatrix} E \\ F \end{pmatrix} = \frac{1}{2i}(Z^* \; I_n)S^t JS\begin{pmatrix} Z \\ I_n \end{pmatrix} = \frac{1}{2i}(Z^* \; I_n)(-J)\begin{pmatrix} Z \\ I_n \end{pmatrix}$$
$$= \frac{1}{2i}(Z^* \; I_n)\begin{pmatrix} O & -I_n \\ I_n & O \end{pmatrix}\begin{pmatrix} Z \\ I_n \end{pmatrix}$$
$$= \frac{1}{2i}(Z - Z^*) > 0 \;.$$

$$(2.1)$$

It follows that,

$$\frac{1}{2i}(E^*F - F^*E) > 0 \;. \tag{2.2}$$

To see F invertible, suppose v is a solution of $Fv = 0$ then we have $v^*F^* = 0$ and

$$v^*(E^*F - F^*E)v = 0 \;.$$

From equation (2.2), we have $v = 0$ which implies F invertible.

Next we show that $\Phi_S(Z)$ is symmetric and $\operatorname{Im}\Phi_S(Z) < 0$. We have

$$(E^t \; F^t)J\begin{pmatrix} E \\ F \end{pmatrix} = (Z^t \; I_n)S^t JS\begin{pmatrix} Z \\ I_n \end{pmatrix} = (Z^t \; I_n)(-J)\begin{pmatrix} Z \\ I_n \end{pmatrix}$$
$$= Z - Z^t = O \;.$$

This implies that $F^t E = E^t F$. From this we get,

$$(EF^{-1})^t = (F^{-1})^t E^t = (F^{-1})^t F^t EF^{-1} = EF^{-1}$$

which means $\Phi_S(Z)$ is symmetric. Moreover we have,

$$\operatorname{Im}\Phi_S(Z) = \operatorname{Im}(EF^{-1}) = \frac{1}{2i}(EF^{-1} - (EF^{-1})^*)$$
$$= \frac{1}{2i}(F^{-1})^*F^*(EF^{-1} - (EF^{-1})^*)FF^{-1}$$
$$= \frac{1}{2i}(F^{-1})^*(F^*E - E^*F)F^{-1} \;.$$

Now for any $v \in \mathbb{C}^n$ and $F^{-1}v \in \mathbb{C}^n$ we have

$$v^* \operatorname{Im}\Phi_S(Z)v = = (F^{-1}v)^*\frac{1}{2i}(F^*E - E^*F)(F^{-1}v).$$

By (2.2), we get $(F^{-1}v)^*\frac{1}{2i}(F^*E - E^*F)(F^{-1}v) < 0$ for any $v \in \mathbb{C}^n$
It follows that $\operatorname{Im}\Phi_S(Z) < 0$. Hence Φ_S maps $\mathcal{S}_n$ to $\overline{\mathcal{S}}_n$.

Finally we show the map Φ_S is onto. Suppose $Z = X + iY \in \overline{\mathcal{S}}_n$. We have

$$S_{Z_-} = \begin{pmatrix} -\sqrt{-Y} & X\sqrt{-Y^{-1}} \\ O & \sqrt{-Y^{-1}} \end{pmatrix} \in A(2n, \mathbb{R})$$

and $\Phi_{S_{Z_-}}(iI_n) = Z$. $-Y$ is a symmetric positive definite matrix, therefore the Spectral Theorem guarantees that $\sqrt{-Y}$ and $\sqrt{-Y^{-1}}$ exist. This completes the proof. $\qquad\square$

As a consequence of this proposition we have the following theorem.

Theorem 2.2. *If $S \in SP_{2n}(\mathbb{C})$ is purely imaginary then Φ_S maps Siegel half space to the space $\overline{\mathcal{S}}_n$.*

Proof. Let $S = iQ$, where

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix} .$$

Since $S^t J S = J$, we have $Q^t J Q = -J$, hence $Q \in A(2n, \mathbb{R})$. Moreover we have

$$\begin{aligned} \Phi_{iQ}(Z) &= (iAZ + iB)(iCZ + iD)^{-1} \\ &= (AZ + B)(CZ + D)^{-1} \\ &= \Phi_Q(Z) . \end{aligned}$$

Since $Q \in A(2n, \mathbb{R})$, by Proposition 2.1 Φ_Q maps $\mathcal{S}_n$ to $\overline{\mathcal{S}}_n$ and hence Φ_{iQ} maps $\mathcal{S}_n$ to $\overline{\mathcal{S}}_n$. Thus completing the proof. $\qquad\square$

In the following theorem we partially classify the matrices from $SP_{2n}(\mathbb{C})$ so that Φ_S is well defined.

Theorem 2.3. *Let*

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in SP_{2n}(\mathbb{C})$$

and $Z \in \mathcal{S}_n$. If $i(S^* J S - J) \geq 0$ then $\Phi_S(Z) = S(Z) = (AZ+B)(CZ+D)^{-1} \in \mathcal{S}_n$.

Proof. For convenience set $E = AZ + B$ and $F = CZ + D$. Suppose $i(S^* J S - J) \geq 0$. For any $Z \in \mathcal{S}_n$ we have

$$(Z^* \ I_n)\,(i(S^* J S - J)) \begin{pmatrix} Z \\ I_n \end{pmatrix} \geq 0 .$$

On the other hand

$$(Z^* \ I_n)\,(i(S^*JS - J)) \begin{pmatrix} Z \\ I_n \end{pmatrix}$$

$$= i\left(\left(S \begin{pmatrix} Z \\ I_n \end{pmatrix} \right)^* \left(S \begin{pmatrix} Z \\ I_n \end{pmatrix} \right) - (Z^* \ I_n) J \begin{pmatrix} Z \\ I_n \end{pmatrix} \right)$$

$$= i(E^* \ F^*) J \begin{pmatrix} F \\ -E \end{pmatrix} - i(Z^* \ I_n) \begin{pmatrix} I_n \\ -Z \end{pmatrix}$$

$$= i(E^*F - F^*E) - \frac{1}{i}(Z - Z^*) \ .$$

Therefore we have $i(E^*F - F^*E) + i(Z - Z^*) \geq 0$ and thus

$$\frac{1}{i}(F^*E - E^*F) \geq \frac{1}{i}(Z - Z^*) = 2\mathrm{Im}(Z) > 0$$

since $Z \in \mathcal{S}_n$. To show that $S(Z) \in \mathcal{S}_n$ we need to show that F is invertible, $S(Z)$ is symmetric, and $\mathrm{Im}(S(Z)) > 0$. First we show that F is invertible. Suppose $Fv = 0$ for $v \in \mathbb{C}^n$, then it follows that $v^*(F^*E - E^*F)v/2i = 0$. Since $-i(F^*E - E^*F) > 0$, we have $v = 0$. Hence F^{-1} exists and EF^{-1} is well-defined.

Next we show symmetry. Notice that since S is a symplectic matrix and $Z \in \mathcal{S}_n$, we have

$$E^tF - F^tE = \begin{pmatrix} Z \\ I_n \end{pmatrix}^t S^t JS \begin{pmatrix} Z \\ I_n \end{pmatrix} = \begin{pmatrix} Z \\ I_n \end{pmatrix}^t J \begin{pmatrix} Z \\ I_n \end{pmatrix} = Z^t - Z = O \ .$$

Therefore $E^tF = F^tE$ and thus by the invertibility of F we have

$$(EF^{-1})^t = (F^{-1})^t E^t = (F^{-1})^t F^t EF^{-1} = EF^{-1} \ ,$$

thus showing the symmetry of $S(Z)$. Finally a similar calculation to the one at the beginning of the proof shows that

$$\mathrm{Im}(S(Z)) = (F^{-1})^* \frac{1}{2i} \left(F^*E - E^*F \right) F^{-1}$$

and since $-i(F^*E - E^*F) > 0$, it follows that $\mathrm{Im}(S(Z)) > 0$. This completes the proof. $\square$

It should be noted the other direction may or may not be true. One of the many issues in showing the other direction is going from $n \times n$ complex matrices to $2n \times 2n$ complex matrices. The direction proven is simpler since the dimensions are going down. In fact, one can give necessary and sufficient conditions on a self-adjoint block matrix when it is positive definite. In fact [1] stated without proof, and [4] proved in unpublished work that is available online, conditions on when a

symmetric block matrix is positive definite. This proof can be easily modified to allow for generic self-adjoint block matrices. We state it without proof.

Proposition 2.4. *Let M be a $2n \times 2n$ self-adjoint block matrix of the form:*

$$M = \begin{pmatrix} \alpha & \beta \\ \beta^* & \gamma \end{pmatrix} \ .$$

The following are equivalent:

(1) $M \geq 0$
(2) $\alpha \geq 0$, $\quad (I_n - \alpha\alpha^\dagger)\beta = O$, $\quad$ *and* $\gamma - \beta^*\alpha^\dagger\beta \geq 0$
(3) $\gamma \geq 0$, $\quad (I_n - \gamma\gamma^\dagger)\beta^* = O$, $\quad$ *and* $\alpha - \beta\gamma^\dagger\beta^* \geq 0$

where $P^\dagger$ is the pseudo-inverse of P.

For more details on pseudo-inverses of square matrices see [5]. Let $M = i(S^*JS - J)$, then it's clear that M is self-adjoint. A computation shows that

$$M = \begin{pmatrix} i(A^*C - C^*A) & i(A^*D - C^*B - I_n) \\ i(B^*C - D^*A + I_n) & i(B^*D - D^*B) \end{pmatrix} := \begin{pmatrix} \alpha & \beta \\ \beta^* & \gamma \end{pmatrix}$$

Based on the discussion from above, we know since S is symplectic, then A^tC and B^tD are symmetric and $A^tD - C^tB = I_n$. Moreover in this case, the entries of S can not be all real, otherwise, we are the case already proven by [3]. Based on this proposition, for $M \geq 0$, we must have

(1) $\quad i(A^*C - C^*A) \geq 0$

(2) $\quad [I_n + (A^*C - C^*A)(A^*C - C^*A)^\dagger](A^*D - C^*B - I_n) = O$

(3) $\quad i(B^*D - D^*B)+$
$\quad\quad + i(B^*C - D^*A + I_n)(A^*C - C^*A)^\dagger(A^*D - C^*B - I_n) \geq 0 \ .$

Whether or not these conditions are equivalent to $i(S^*JS - J) \geq 0$ remains unknown.

As discussed in Example 1.1 and Example 1.2, not all matrices from $SP_{2n}(\mathbb{C})$ give a well defined map $\Phi_S(Z)$ for $Z \in \mathcal{S}_n$ nor does the map always preserve $\mathcal{S}_n$. Define the following set:

$\Omega = \{S \in SP_{2n}(\mathbb{C}) : \Phi_S : \mathcal{S}_n \to \mathcal{S}_n \text{ is well defined}\}$, so that we can now consider a map $\tilde{\Phi} : \Omega \times \mathcal{S}_n \to \mathcal{S}_n$ which will be an extension of Φ defined in equation (1.3). The next natural question that arises is: does Ω form a subgroup of $SP_{2n}(\mathbb{C})$? That way we could consider $\tilde{\Phi}$ as

a group action of Ω on $\mathcal{S}_n$. Unfortunately the following example shows that this does not happen.

Example 2.5. *Let* $S = \begin{pmatrix} I_n & iI_n \\ O & I_n \end{pmatrix}$ *then* $S \in SP_{2n}(\mathbb{C})$ *and* Φ_S *is clearly well defined as* $\Phi_S(Z) = Z + iI \in \mathcal{S}_n$ *for all* $Z \in \mathcal{S}_n$. *However, for* $S^{-1} = \begin{pmatrix} I_n & -iI_n \\ O & I_n \end{pmatrix}$, $\Phi_{S^{-1}}(iI_n) = 0 \notin \mathcal{S}_n$

Even though this example shows Ω is not a subgroup of $SP_{2n}(\mathbb{C})$, the next proposition shows that Ω is at least closed under matrix multiplication.

Proposition 2.6. *Let* $S, R \in \Omega$. *Then* $\Phi_S \circ \Phi_R = \Phi_{SR}$ *and* $SR \in \Omega$.

Proof. Let $S, R \in \Omega$ have the following form: $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ and $R = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$. Since $S, R \in \Omega$, for any $Z \in \mathcal{S}_n$, we have that $\Phi_S(Z)$ and $\Phi_R(Z)$ are well defined, moreover $\Phi_R(Z) \in \mathcal{S}_n$. Thus the composition $\Phi_S \circ \Phi_R$ makes sense and hence we have

$$
\begin{aligned}
\Phi_S \circ \Phi_R(Z) &= \Phi_S(\Phi_R(Z)) = \Phi_S((EZ + F)(GZ + H)^{-1}) \\
&= \left(A(EZ + F)(GZ + H)^{-1} + B\right) \\
&\quad \times \left(C(EZ + F)(GZ + H)^{-1} + D\right)^{-1} \\
&= \left(A(EZ + F)(GZ + H)^{-1} + B\right)(GZ + H) \\
&\quad \times (GZ + H)^{-1}\left(C(EZ + F)(GZ + H)^{-1} + D\right)^{-1} \\
&= (A(EZ + F) + B(GZ + H)) \\
&\quad \times (C(EZ + F) + D(GZ + H))^{-1} \\
&= ((AE + BG)Z + (AF + BH)) \\
&\quad \times ((CE + DG)Z + (CF + DH)))^{-1} \\
&= \Phi_{SR}(Z) \ .
\end{aligned}
$$

Next since $SP_{2n}(\mathbb{C})$ is a group, $SR \in SP_{2n}(\mathbb{C})$. The above composition works for an arbitrary $Z \in \mathcal{S}_n$, thus $\Phi_{SR}(Z)$ is well defined. Finally since $\Phi_S, \Phi_R : \mathcal{S}_n \to \mathcal{S}_n$, the above composition shows $\Phi_{SR} : \mathcal{S}_n \to \mathcal{S}_n$. Hence $SR \in \Omega$ completing the proof. $\square$

Again [3, 8] showed that the map $\Phi \colon SP_{2n}(\mathbb{R}) \times \mathcal{S}_n \to \mathcal{S}_n$, which was defined in equation (1.3), forms a group action. In addition, we have the following proposition that was also due to [3].

Proposition 2.7. $\Phi : SP_{2n}(\mathbb{R}) \times \mathcal{S}_n \to \mathcal{S}_n$ *is a transitive action.*

Proof. We show that $\mathcal{S}_n$ has a single orbit. In other words, every $Z \in \mathcal{S}_n$ is an image of some $S \in SP_{2n}(\mathbb{R})$. Let $Z = X + iY \in \mathcal{S}_n$. Define S_Z by

$$S_Z = \begin{pmatrix} \sqrt{Y} & X\sqrt{Y^{-1}} \\ O & \sqrt{Y^{-1}} \end{pmatrix} .$$

Recall Y is a symmetric positive definite matrix, therefore the Spectral Theorem guarantees that $\sqrt{Y}$ and $\sqrt{Y^{-1}}$ exist. It is easy to verify that $(S_Z)^t J S_Z = J$. Therefore, $S_Z \in SP_{2n}(\mathbb{R})$. Now consider $S_Z(iI_n)$. Indeed we have

$$S_Z(iI_n) = (\sqrt{Y}iI_n + X\sqrt{Y^{-1}})(\sqrt{Y^{-1}})^{-1} = X + iY = Z$$

thus showing $\mathcal{S}_n$ has a single orbit completing the proof. $\qquad\square$

Since the action of

$$I_{2n} := \begin{pmatrix} I_n & O \\ O & I_n \end{pmatrix}$$

is same as the action of

$$-I_{2n} = \begin{pmatrix} -I_n & O \\ O & -I_n \end{pmatrix}$$

we can identify them so that we can form the quotient space

$$SP_{2n}(\mathbb{R})/\{I_{2n}, -I_{2n}\}$$

again denoted by $SP_{2n}(\mathbb{R})$. This makes the action of $SP_{2n}(\mathbb{R})$ on $\mathcal{S}_n$ bijective. The following proposition was stated in [2] again without proof. Like before we present a proof for completeness of this paper.

Proposition 2.8. *The set of matrices*

$$K = \{U \in SP_{2n}(\mathbb{R}) : \Phi_U(iI_n) = iI_n\}$$

is precisely the subgroup of orthogonal matrices in $SP_{2n}(\mathbb{R})$.

Proof. Suppose

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in SP_{2n}(\mathbb{R})$$

such that $\Phi_U(iI_n) = iI_n$ then $(AiI_n + B)(CiI_n + D)^{-1} = iI_n$. This implies that $iA + B = iI_n(iC + D)$, therefore it follows that $B = -C$ and $A = D$. Hence U and U^t are of the forms

$$U = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \text{ and } U^t = \begin{pmatrix} A^t & -B^t \\ B^t & A^t \end{pmatrix} .$$

Thus we have

$$UU^t = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \begin{pmatrix} A^t & -B^t \\ B^t & A^t \end{pmatrix}$$
$$= \begin{pmatrix} AA^t + BB^t & -AB^t + BA^t \\ -BA^t + AB^t & BB^t + AA^t \end{pmatrix} = I_{2n}$$

since $AA^t + BB^t = I_n$ and $-AB^t$ is symmetric. This shows that $U^t = U^{-1}$, hence U is orthogonal. Moreover if

$$U = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in K ,$$

then $U^{-1} \in K$. This is true since

$$\Phi_{U^{-1}}(iI_n) = \begin{pmatrix} A^t & -B^t \\ B^t & A^t \end{pmatrix}(iI_n) = \left(A^t(iI_n) - B^t \right) \left(B^t(iI_n) + A^t \right)^{-1}$$
$$= (Ai - B)^t \left[(B(iI_n) + A)^t \right]^{-1} = \left[(Bi + A)^{-1}(Ai - B) \right]^t$$
$$= \left[i(Bi + A)^{-1}(A + Bi) \right]^t = iI_n .$$

In addition, if U and V are both in K then by Proposition 2.6 we have $\Phi_{UV^{-1}}(iI_n) = \Phi_U(\Phi_{V^{-1}}(iI_n)) = \Phi_U(iI_n)) = iI_n$. This shows that K is a subgroup. Thus the result follows. $\square$

Since K is the subgroup of orthogonal matrices in $SP_{2n}(\mathbb{R})$ general theory tells us that K is a normal subgroup. Therefore we can consider the quotient space

$$SP_{2n}(\mathbb{R})/K = \{SK : S \in SP_{2n}(\mathbb{R})\}$$

where $SK = \{SU : U \in K\}$ is a coset of K. As shown in [3], this space is metrizable with a metric given by

$$d^S(S_1K, S_2K) = 2\ln \|S_1^{-1}S_2\| .$$

Notice that the metric is independent of the choice of representative in the equivalence class SK.

Define a map Ψ in the following way

$$\Psi : SP_{2n}(\mathbb{R})/K \to \mathcal{S}_n \text{ via } SK \mapsto \Phi_S(iI_n) . \tag{2.3}$$

Since Φ_S, when acting $SP_{2n}(\mathbb{R})$, is a bijective map, the map Ψ is also bijective. The following theorem was proved in [2].

Theorem 2.9. *The map Ψ, defined in equation (2.3), is an isometry for the metrics d^S and d_∞, that is*

$$d^S(S_1K, S_2K) = d_\infty(S_1(iI_n), S_2(iI_n)) .$$

In particular,

$$d_\infty(Z_1, Z_2) = 2\ln \|S_{Z_1}^{-1}S_{Z_2}\| .$$

We skip the proof for brevity. Based on this result we have the following theorem.

Theorem 2.10. *If $S \in SP_{2n}(\mathbb{C})$ is purely imaginary then the map Φ_S is an isometry between the metric spaces $(\mathcal{S}_n, d_\infty)$ and $(\overline{\mathcal{S}}_n, d_-)$.*

Proof. Let $S = iQ$, where

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix} .$$

Since $S^t J S = J$ we have $Q^t J Q = -J$, hence $Q \in A(2n, \mathbb{R})$. For any $Z = X + iY \in \mathcal{S}_n$ we have $\Phi_{S_Z}(iI_n) = S_Z(iI_n) = Z$. Let

$$S_{Z-} = \begin{pmatrix} -\sqrt{Y} & X\sqrt{Y^{-1}} \\ O & \sqrt{Y^{-1}} \end{pmatrix} \text{ and } I_- = \begin{pmatrix} -I_n & O \\ O & I_n \end{pmatrix} .$$

Then $S_{Z-} = S_Z I_-$ and $S_{Z-}^t J S_{Z-} = -J$ implying that $S_{Z-} \in A(2n, \mathbb{R})$. Since both Q and S_{Z-} are in $A(2n, \mathbb{R})$ this implies that QS_{Z-} is in $SP_{2n}(\mathbb{R})$. Notice that $S_{Z-}(iI_n) = \overline{Z}$. Therefore we have

$$\begin{aligned}
d_-(\Phi_S(Z), \Phi_S(W)) &= d_-(\Phi_{iQ}(Z), \Phi_{iQ}(W) = d_-(\Phi_Q(Z), \Phi_Q(W)) \\
&= d_\infty(\overline{\Phi_Q(Z)}, \overline{\Phi_Q(W)}) = d_\infty(\Phi_Q(\overline{Z}), \Phi_Q(\overline{W})) \\
&= d_\infty(QS_{Z-}(iI_n), QS_{W-}(iI_n)) \\
&= 2\ln \|(QS_{Z-})^{-1}QS_{W-}\| = 2\ln \|S_{Z-}^{-1}S_{W-}\| \\
&= 2\ln \|I_- S_Z^{-1} S_W I_-\| = 2\ln \|S_Z^{-1}S_W\| \\
&= d_\infty(Z, W)
\end{aligned}$$

since I_- is a unitary. Thus this completes the proof. $\qquad\square$

Unfortunately in general, the map Φ_S is not an isometry. In fact we have the following theorem.

Theorem 2.11. *If the matrix S is symplectically similar to*

$$R = \begin{pmatrix} I_n & Z \\ I_n & 0 \end{pmatrix} \in SP_{2n}(\mathbb{C})$$

where $Z \in \mathcal{S}_n$, that is $S = PRP^{-1}$ for some $P \in SP_{2n}(\mathbb{R})$ then there exists a δ with $0 < \delta < 1$ such that for all $Z_1, Z_2 \in \mathcal{S}_n$,

$$d_\infty(S(Z_1), S(Z_2)) \leq \delta d_\infty(Z_1, Z_2) .$$

Proof. Claim: for $Z = X + iY, W = W_1 + iW_2 \in \mathcal{S}_n$

$$F_{W+Z}(\gamma) < \delta F_W(\gamma)$$

for some $0 < \delta < 1$. Indeed we have

$$F_{W+Z}(\gamma) = F_{(W_1+X)+i(W_2+Y)}(\gamma) = \|(W_2 + Y)^{-\frac{1}{2}}\gamma(W_2 + Y)^{-\frac{1}{2}}\|$$

$$= \|(W_2 + Y)^{-\frac{1}{2}}W_2^{\frac{1}{2}}W_2^{-\frac{1}{2}}\gamma W_2^{-\frac{1}{2}}W_2^{\frac{1}{2}}(W_2 + Y)^{-\frac{1}{2}}\|$$

$$\leq \|(W_2 + Y)^{-\frac{1}{2}}W_2^{\frac{1}{2}}\|^2 \|(W_2^{-\frac{1}{2}}\gamma W_2^{-\frac{1}{2}}\| .$$

Since $Y > 0$, we have $W_2^{\frac{1}{2}} < (W_2 + Y)^{\frac{1}{2}}$ and $(W_2 + Y)^{-\frac{1}{2}} > 0$. It follows

$$(W_2 + Y)^{-\frac{1}{2}}W_2^{\frac{1}{2}} < (W_2 + Y)^{-\frac{1}{2}}(W_2 + Y)^{\frac{1}{2}} = I_n .$$

Therefore we have $\|(W_2 + Y)^{-\frac{1}{2}}W_2^{\frac{1}{2}}\| < 1$. Thus we can choose

$$\delta = \|(W_2 + Y)^{-\frac{1}{2}}W_2^{\frac{1}{2}}\|$$

proving the claim. Now let $\gamma(t)$ be a path joining Z_1, Z_2 so that $\gamma(t) + Z$ is a path joining $Z_1 + Z, Z_2 + Z$ and

$$\int_0^1 F_{\gamma(t)+Z}(\dot{\gamma}(t))\, dt < \delta \int_0^1 F_{\gamma(t)}(\dot{\gamma}(t))\, dt .$$

Taking the infimum over all the paths $\gamma(t)$ joining Z_1 and Z_2, we get

$$\int_0^1 F_{\gamma(t)+Z}(\dot{\gamma}(t))\, dt < \delta d_\infty(Z_1, Z_2) .$$

Again taking the infimum over all the paths $\gamma(t) + Z$ joining $Z_1 + Z$ and $Z_2 + Z$, we get

$$d_\infty(R(Z_1), R(Z_2)) \leq \delta d_\infty(Z_1, Z_2) .$$

Finally we get

$$\begin{aligned}
d_\infty(S(Z_1), S(Z_2)) &= d_\infty(PRP^{-1}(Z_1), PRP^{-1}(Z_2)) \\
&= d_\infty(RP^{-1}(Z_1), RP^{-1}(Z_2)) \\
&\leq \delta d_\infty(P^{-1}(Z_1), P^{-1}(Z_2)) \\
&= \delta d_\infty(Z_1, Z_2) ,
\end{aligned}$$

completing the proof. □

Recall that the Finsler metric $d_\infty(\cdot, \cdot)$ becomes the hyperbolic metric when $n = 1$. Therefore we can compare distances of numbers in the upper half plane and corresponding counterparts in the Siegel space. We end the paper with the following result.

Theorem 2.12. *Let $\mathbb{B}(0)$ be the unit ball in $\mathbb{C}^n$. For any $Z_1, Z_2 \in \mathcal{S}_n$ and $v \in \mathbb{B}(0)$,*

$$d_\infty(v^* Z_1 v, v^* Z_2 v) \le d_\infty(Z_1, Z_2) \ .$$

Proof. If $Z(t)$ is a path joining Z_1 and Z_2 then $v^* Z(t) v$ is a path joining $v^* Z_1 v$ and $v^* Z_2 v$ with $v \in \mathbb{B}(0)$. For $Z(t) = X(t) + iY(t)$ and $v \in \mathbb{B}(0)$ we have

$$\begin{aligned}
\text{Im}(v^* Z(t) v) &= \frac{1}{2i}(v^* Z(t) v - (v^* Z(t) v)^*) \\
&= \frac{1}{2i} v^* (Z(t) - Z(t)^*) v^* \\
&= v^* Y(t) v^*
\end{aligned}$$

with $v^* Y(t) v > 0$. Recall that for a $n \times n$ matrix A and a vector x we have the following submultiplicative estimate: $\|Ax\| \le \|A\|\|x\|$. Using this, the Cauchy-Schwarz inequality, and the fact that the Finsler metric becomes the hyperbolic metric in one dimension we get

$$\begin{aligned}
F_{v^* Z(t) v}(v^* \dot{Z}(t) v) &= \|(v^* Y v)^{-\frac{1}{2}} (v^* \dot{Z}(t) v)(v^* Y v)^{-\frac{1}{2}}\| \\
&= \frac{1}{v^* Y v} \|v^* Y^{\frac{1}{2}} Y^{-\frac{1}{2}} \dot{Z}(t) Y^{-\frac{1}{2}} Y^{\frac{1}{2}} v\| \\
&= \frac{1}{v^* Y v} \langle Y^{\frac{1}{2}} v, Y^{-\frac{1}{2}} \dot{Z}(t) Y^{-\frac{1}{2}} Y^{\frac{1}{2}} v \rangle \\
&\le \frac{1}{v^* Y v} \|Y^{\frac{1}{2}} v\|^2 \|Y^{-\frac{1}{2}} \dot{Z}(t) Y^{-\frac{1}{2}}\| \\
&= \|Y^{-\frac{1}{2}} \dot{Z}(t) Y^{-\frac{1}{2}}\| = F_{Z(t)}(\dot{Z}(t)) \ .
\end{aligned}$$

Upon integration we get,

$$\int_0^1 F_{v^* Z(t) v}(v^* \dot{Z}(t) v) \ dt \le \int_0^1 F_{Z(t)}(\dot{Z}(t)) \ dt \ .$$

As before, taking the infimum over all the paths $Z(t)$ joining Z_1 and Z_2, we get

$$\int_0^1 F_{v^* Z(t) v}(v^* \dot{Z}(t) v) \ dt \le d_\infty(Z_1, Z_2) \ .$$

Finally taking the infimum over all the paths $v^*Z(t)v$ joining v^*Z_1v and v^*Z_2v, we get

$$d_\infty(v^*Z_1v, v^*Z_2v) \leq d_\infty(Z_1, Z_2) \ .$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[1] Boyd, S. and Vandenberghe, L., *Convex Optimization*, Cambridge University Press, 2004.

[2] Froese, R., Hasler, D. and Spitzer, W., Transfer Matrices, Hyperbolic Geometry and Absolutely Continuous Spectrum for Some Discrete Schrödinger Operators on Graphs, *Jour. Func. Anal.*, 230(1), 184-221, 2006.

[3] Freitas, P. J., On the Action of the Symplectic Group on the Siegel Upper Half Plane, PhD thesis, University of Illinois at Chicago, 1999.

[4] Gallier, J., The Schur Complement and Symmetric Positive Semidefinite (and Definite) Matrices, unpublished work, available online http://www.cis.upenn.edu/ jean/schur-comp.pdf.

[5] Gallier, J. *Geometric Methods and Applications, For Computer Science and Engineering*, TAM, Vol. 38, Springer, 2000.

[6] Ohsawa, T., The Siegel Upper Half Space is a Marsden-Weinstein Quotient: Symplectic Reduction and Gaussian Wave Packets, *Lett. Math. Phys.*, 105(9), 1301-1320, 2015.

[7] Pontoja, J., Andrade, J. S. and Vargas, J. A., On the Construction of a Finite Siegel Space, *Jour. of Lie Theo.*, 25, 1045-1071, 2015.

[8] Siegel, C. L., *Symplectic Geometry*, Academic Press. New York and London, 1964.

Department of Mathematics, Embry-Riddle Aeronautical University, 600 S. Clyde Moris Blvd., Daytona Beach, FL 32114, U.S.A.
  Email address: `acharyak@erau.edu`

Department of Mathematics and Statistics, Mississippi State University, 175 President's Cir., Mississippi State, MS 39762, U.S.A.
  Email address: `mmcbride@math.msstate.edu`