GENERALIZATION OF THE HILBERT METRIC TO THE SPACE OF POSITIVE DEFINITE MATRICES

Carlangelo Liverani
Maciej P. Wojtkowski

Abstract. We introduce a generalization of the Hilbert projective metric to the space of positive definite matrices which we view as part of the Lagrangian Grassmannian.

§0. Introduction.

In his treatment of Kalman Bucy filters Bougerol [B1], [B2] uses the Riemannian metric on the set of positive definite matrices considered as a Riemannian symmetric space.

Graphs of symmetric linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ are Lagrangian subspaces in the standard linear symplectic space $\mathbb{R}^n \times \mathbb{R}^n$. We call a Lagrangian subspace positive if it is a graph of a positive definite linear map. Further we call a linear symplectic map monotone if it maps positive Lagrangian subspaces onto positive Lagrangian subspaces. Bougerol discovered that the symplectic matrices in Kalman filtering theory are monotone. He shows that the action of any monotone map on the manifold of positive Lagrangian subspaces contracts the metric of the Riemannian symmetric space. It is the only (up to scale) Riemannian metric which has this property.

The goal of this paper is to introduce a natural Finsler metric in the manifold of positive definite matrices which in addition to being contracted by the action of any monotone map has striking geometric properties. In particular we obtain that the coefficient of least contraction is equal to the hyperbolic tangent of one half of the diameter of the image. This is the same relation which was obtained by Birkhoff [Bir] for the Hilbert projective metric.

In the case of the positive orthant the Hilbert metric is also only Finsler (cf. [W2]) which reflects the nonsmoothness of the cone. It is natural that the generalization of the Hilbert metric to the space of positive definite matrices is not smooth because its boundary in the Lagrangian Grassmannian is not smooth.

After this paper was written we learned that this metric was discussed earlier by Vesentini [Ves] from a completely different point of view which is applicable also

The second author gratefully acknowledges the hospitality of Forschungsinstitut für Mathematik at ETH Zürich, where this paper was written. He is partially supported by the NSF Grant DMS–9017993 and the Sloan Foundation

in the infinite dimensional setting. In Vesentini's approach our metric becomes a Caratheodory-type metric on the convex cone of positive definite matrices.

§1. Preliminaries.

We consider a linear symplectic space $\mathcal{W}$ of dimension $2n$ with the symplectic form ω. We call $\mathcal{W} = \mathbb{R}^n \times \mathbb{R}^n$ the standard linear symplectic space if

$$\omega(w_1, w_2) = \langle \xi^1, \eta^2 \rangle - \langle \xi^2, \eta^1 \rangle,$$

$$\text{where } w_i = (\xi^i, \eta^i), \ i = 1, 2, \quad \text{and} \quad \langle \xi, \eta \rangle = \xi_1 \eta_1 + \cdots + \xi_n \eta_n.$$

The symplectic group $Sp\,(n, \mathbb{R})$ is the group of linear maps of $\mathcal{W}$ ($2n \times 2n$ matrices if $\mathcal{W} = \mathbb{R}^n \times \mathbb{R}^n$) preserving the symplectic form i.e., $L \in Sp\,(n, \mathbb{R})$ if

$$\omega(Lw_1, Lw_2) = \omega(w_1, w_2)$$

for every $w_1, w_2 \in \mathcal{W}$.

By definition a Lagrangian subspace of a linear symplectic space $\mathcal{W}$ is an n-dimensional subspace on which the restriction of ω is zero (equivalently it is a maximal subspace on which ω vanishes).

Definition 1. *Given two transversal Lagrangian subspaces V_1 and V_2 we define the sector between V_1 and V_2 by*

$$\mathcal{C} = \mathcal{C}\,(V_1, V_2) = \{w \in \mathcal{W} | \omega(v_1, v_2) \geq 0 \textit{ for } w = v_1 + v_2, v_i \in V_i, i = 1, 2\}.$$

We call V_1 and V_2 the sides of the sector.

Equivalently we define first the quadratic form associated with an ordered pair of transversal Lagrangian subspaces

$$\mathcal{Q}(w) = 2\omega(v_1, v_2)$$

where $w = v_1 + v_2, v_i \in V_i, i = 1, 2,$ is the unique decomposition. The factor 2 is introduced here to simplify some of the future formulas. We have now

$$\mathcal{C} = \{w \in \mathcal{W} | \mathcal{Q}(w) \geq 0\}.$$

In the standard symplectic space if we take $V_1 = \mathbb{R}^n \times \{0\}$ and $V_2 = \{0\} \times \mathbb{R}^n$ we get

$$\mathcal{Q}\,((\xi, \eta)) = 2\langle \xi, \eta \rangle$$

and

$$\mathcal{C} = \{(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^n | \langle \xi, \eta \rangle \geq 0\}.$$

Since any two pairs of transversal Lagrangian subspaces are symplectically equivalent we can consider only this case without any loss of generality. We will keep using the coordinate free language to expose the geometric character of some formulas.

The Lagrangian subspaces which are contained in the boundary of the sector $\mathcal{C}$ can be described in the following way. Let $Y_1 \subset V_1$ be any linear subspace and let $Y_2 \subset V_2$ be the intersection of V_2 with the skeworthogonal complement of Y_1. Then $Y_1 + Y_2$ is a Lagrangian subspace in the boundary of the sector. It can be shown [W4] that all Lagrangian subspaces in the boundary of the sector are obtained in this way. We see that the subspaces V_1 and V_2 are singled out among all the Lagrangian subspaces contained in the boundary of $\mathcal{C}$ by being the only isolated points. It follows that a sector determines uniquely its sides.

Based on the notion of the sector between two transversal Lagrangian subspaces (or the quadratic form $\mathcal{Q}$) we define two monotonicity properties of a linear symplectic map.

Definition 2. *Given the sector $\mathcal{C}$ between two transversal Lagrangian subspaces we call a linear symplectic map L monotone if*

$$L\mathcal{C} \subset \mathcal{C}$$

and strictly monotone if

$$L\mathcal{C} \setminus \{0\} \subset int\mathcal{C}.$$

Equivalent property is given by

Theorem 1. *L is (strictly) monotone if and only if $\mathcal{Q}\,(Lw) \geq \mathcal{Q}\,(w)$ for every $w \in \mathcal{W}$ ($\mathcal{Q}\,(Lw) > \mathcal{Q}\,(w)$ for every $w \in \mathcal{W},\ w \neq 0$).*

The increasing of the quadratic form implies trivially monotonicity but the converse holds only because of a very special geometric structure of the sector and it is not true in general for cones defined by arbitrary quadratic forms. This theorem was first proved in [W1].

Let

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

be a symplectic map of the standard symplectic space $\mathbb{R}^n \times \mathbb{R}^n$ monotone with respect to the standard sector. A, B, C, D are $n \times n$ matrices and it can be shown that A and D are nondegenerate (cf.[W3],[W4]).

Let us describe those symplectic matrices which are monotone in the weakest form, namely they preserve the quadratic form $\mathcal{Q}$. We will call such matrices $\mathcal{Q}$-isometries.

Proposition 1. *If L is a linear symplectic map and*

$$L\mathcal{C} = \mathcal{C}$$

then

$$L = \begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix}.$$

In particular it preserves the quadratic form $\mathcal{Q}$

$$\mathcal{Q} \circ L = \mathcal{Q}.$$

Proof. If $L\mathcal{C} = \mathcal{C}$ then L maps also the boundary of the sector $\mathcal{C}$ onto itself. Since the sides of the sector are the only isolated Lagrangian subspaces in the boundary of the sector it follows that they stay put under L. Hence $B = C = 0$. By symplecticity $D = A^{*-1}$. $\square$

Given a monotone L we can always factor out the following $\mathcal{Q}$-isometry on the left.

$$L = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix} \begin{pmatrix} I & \cdot \\ \cdot & \cdot \end{pmatrix}.$$

Symplecticity of L forces further unique factorization

$$L = \begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ P & I \end{pmatrix} \begin{pmatrix} I & R \\ 0 & I \end{pmatrix}$$

with symmetric P and R. Moreover monotonicity forces $P \geq 0$ and $R \geq 0$. Strict monotonicity means that $P > 0$ and $R > 0$. These claims follow from the following

Proof of Theorem 1. Using the above factorization we get for $w = (\xi, \eta)$

$$\mathcal{Q}(Lw) = \langle \xi, \eta \rangle + \langle R\eta, \eta \rangle + \langle P(\xi + R\eta), \xi + R\eta \rangle.$$

Putting $\eta = 0$ we obtain that $P \geq 0$. To show that also $R \geq 0$ let us consider an eigenvector η of R with eigenvalue λ and let $w = (a\eta, \eta)$. We get that if $a \geq 0$ then $w \in \mathcal{C}$ and $Lw \in \mathcal{C}$ so that

$$(a + \lambda)\langle \eta, \eta \rangle + (a + \lambda)^2 \langle P\eta, \eta \rangle \geq 0.$$

This implies immediately that $\lambda \geq 0$ which ends the proof of the monotone version of the Theorem. The strictly monotone version is obtained in a similar way. $\square$

As a byproduct of the proof we get the following useful observation

Proposition 2. *A monotone map L is strictly monotone if and only if*

$$LV_i \setminus \{0\} \subset int\mathcal{C}, \ \ i = 1, 2.$$

$\square$

§2. The metric in the manifold of positive Lagrangian subspaces.

We start by defining the symplectic angle $\Theta(u, w)$ between two vectors in the interior of the sector $\mathcal{C}$

Definition 3. *The symplectic angle between two vectors $u, w \in int\mathcal{C}$ is the real number $\Theta(u, w)$ such that*

$$\omega(u, w) = \sqrt{\mathcal{Q}(u)}\sqrt{\mathcal{Q}(w)} \sinh \Theta(u, w).$$

For any monotone symplectic map L and two vectors $u,\, w \in \mathrm{int}\mathcal{C}$ we have

$$\sinh\Theta(Lu,\, Lw) = \sinh\Theta(u,\, w)\sqrt{\frac{\mathcal{Q}(u)}{\mathcal{Q}(Lu)}}\sqrt{\frac{\mathcal{Q}(w)}{\mathcal{Q}(Lw)}}. \tag{1}$$

Hence by Theorem 1 a monotone map does not increase the absolute value of the symplectic angle and the contraction of the symplectic angle can be estimated by the least coefficient of expansion of $\mathcal{Q}$ under the monotone map. We will find an explicit formula for this coefficient. For a linear symplectic map L monotone with respect to the sector $\mathcal{C}$ we define the coefficient of expansion at $w \in \mathrm{int}\mathcal{C}$ by

$$\beta\,(w, L) = \sqrt{\frac{\mathcal{Q}\,(Lw)}{\mathcal{Q}\,(w)}}.$$

We define further the least coefficient of expansion $\sigma\,(L)$ by

$$\sigma\,(L) = \inf_{w\in\mathrm{int}\mathcal{C}} \beta\,(w, L)\,.$$

To find the value of this expansion coefficient we will use the fact that it does not change if L is multiplied on the left or on the right by $\mathcal{Q}$-isometries.

Proposition 3. *If*

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

is a strictly monotone map then by multiplying it by $\mathcal{Q}$-isometries on the left and on the right we can bring it to the form

$$\begin{pmatrix} I & I \\ T & I+T \end{pmatrix}$$

where T is diagonal and has the same eigenvalues as C^*B.

Proof. The factorization of a monotone map yields

$$\begin{pmatrix} A & 0 \\ 0 & A^{*-1} \end{pmatrix} L = \begin{pmatrix} I & R \\ P & I+PR \end{pmatrix}$$

where $P > 0$, $R > 0$ and $PR = C^*B$.

We have further

$$\begin{pmatrix} R^{-\frac{1}{2}} & 0 \\ 0 & R^{\frac{1}{2}} \end{pmatrix}\begin{pmatrix} I & R \\ P & I+PR \end{pmatrix}\begin{pmatrix} R^{\frac{1}{2}} & 0 \\ 0 & R^{-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} I & I \\ K & I+K \end{pmatrix}$$

where $K = R^{\frac{1}{2}}PR^{\frac{1}{2}}$ has the same eigenvalues as $C^*B = PR$.

Finally if F is the orthogonal matrix which diagonalizes K, i.e., $F^{-1}KF$ is diagonal, then

$$\begin{pmatrix} F^{-1} & 0 \\ 0 & F^{-1} \end{pmatrix}\begin{pmatrix} I & I \\ K & I+K \end{pmatrix}\begin{pmatrix} F & 0 \\ 0 & F \end{pmatrix} = \begin{pmatrix} I & I \\ T & I+T \end{pmatrix}$$

has the desired form with $T = F^{-1}KF$ having the same eigenvalues as C^*B. $\square$

For a monotone L the matrix C^*B is equal to the product of two positive semi-definite matrices ($C^*B = PR$) and so it has only real nonnegative eigenvalues. Let us denote them by $0 \le t_1 \le \cdots \le t_n$.

Proposition 4. *For a monotone map L*

$$\sigma(L) = \sqrt{1+t_1} + \sqrt{t_1} = \exp \sinh^{-1} \sqrt{t_1}$$

and if L is strictly monotone

$$\sigma(L) = \beta(w, L)$$

for some $w \in int\mathcal{C}$.

Proof. Let us put

$$m(L) = \sqrt{1+t_1} + \sqrt{t_1} = \min_{1 \le i \le n} \left(\sqrt{1+t_i} + \sqrt{t_i} \right).$$

First we prove the inequality $\beta(w, L) \geq m(L)$ for $w \in \mathrm{int}\mathcal{C}$. Since both $\beta(w, L)$ and $m(L)$ are continuous functions of L it is sufficient to prove the inequality for strictly monotone maps only. In view of Proposition 3 we can take

$$L = \begin{pmatrix} I & I \\ T & I+T \end{pmatrix}$$

with diagonal T and $t_1, \ldots, t_n$ on the diagonal. We compute directly for $w = (\xi, \eta)$ such that $\mathcal{Q}(w) = 1$

$$\begin{aligned}
(\beta(w, L))^2 &= \sum_{i=1}^{n} \left(t_i \xi_i^2 + (1 + 2t_i)\, \xi_i \eta_i + (1 + t_i)\, \eta_i^2 \right) \\
&= \sum_{i:\xi_i\eta_i \geq 0} \left(\left(\sqrt{t_i}\xi_i - \sqrt{1+t_i}\eta_i \right)^2 + \left(\sqrt{1+t_i} + \sqrt{t_i} \right)^2 \xi_i \eta_i \right) \\
&+ \sum_{i:\xi_i\eta_i \langle 0} \left(\left(\sqrt{t_i}\xi_i + \sqrt{1+t_i}\eta_i \right)^2 + \left(\sqrt{1+t_i} - \sqrt{t_i} \right)^2 \xi_i \eta_i \right) \geq \\
&\geq \sum_{i:\xi_i\eta_i \geq 0} \left(\sqrt{1+t_i} + \sqrt{t_i} \right)^2 \xi_i \eta_i + \sum_{i:\xi_i\eta_i < 0} \left(\sqrt{1+t_i} + \sqrt{t_i} \right)^{-2} \xi_i \eta_i \geq \\
&\geq (1 + \delta)\,(m(L))^2 - \delta\,(m(L))^{-2} \geq (m(L))^2
\end{aligned}$$

where

$$\delta = \left(\sum_{i:\xi_i\eta_i \geq 0} \xi_i \eta_i \right) - 1 \geq 0$$

and all the inequalities become equalities with appropriate choice of w.

Thus the Proposition is proven for strictly monotone matrices. To extend it to all monotone matrices we proceed as follows. For any $\epsilon > 0$ we choose a strictly monotone matrix L_ϵ so close to the identity that $m(L_\epsilon L) < m(L) + \epsilon$. Let $w_\epsilon \in \mathrm{int}\mathcal{C}$ be such that

$$\beta(w_\epsilon, L_\epsilon L) = m(L_\epsilon L) = \sigma(L_\epsilon L).$$

But $\beta(w, L_\epsilon L) > \beta(w, L)$ for any $w \in \mathrm{int}\mathcal{C}$. Hence

$$m(L) \leq \sigma(L) \leq \beta(w_\epsilon, L) < \beta(w_\epsilon, L_\epsilon L) = m(L_\epsilon L) < m(L) + \epsilon$$

which ends the proof. $\square$

We say that a Lagrangian subspace is contained strictly in the sector $\mathcal{C}$ or that it is positive (if it is clear which sector we have in mind) if all of its nonzero vectors are contained in the interior of $\mathcal{C}$. We consider the manifold of all Lagrangian subspaces which are strictly contained in the sector $\mathcal{C}$ and denote it by $\mathcal{L}$. Although $\mathcal{C}$ is a smooth (nonconvex) cone the boundary of $\mathcal{L}$ in the Lagrangian Grassmannian is not smooth.

We define further the distance of two Lagrangian subspaces U, W contained strictly in the sector $\mathcal{C}$.

Definition 4. *The distance $s(U, W)$ of two positive Lagrangian subspaces U and W is equal to the supremum of absolute values of symplectic angles between nonzero vectors from the two Lagrangian subspaces, i.e.,*

$$s(U, W) = \sup_{\substack{0 \neq u \in U \\ 0 \neq w \in W}} |\Theta(u, w)|.$$

We will check that $s(U, W)$ is indeed a metric by computing it in coordinates. We assume that we have the standard linear symplectic space $\mathbb{R}^n \times \mathbb{R}^n$ and $V_1 = \mathbb{R}^n \times \{0\}$ and $V_2 = \{0\} \times \mathbb{R}^n$. Let $U : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map and

$$gU = \{(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^n \,|\, \eta = U\xi\}$$

be its graph. The subspace gU is Lagrangian if and only if U is symmetric. Moreover for a symmetric U the subspace $gU \subset \mathcal{C}$ if and only if $U \geq 0$. For a monotone L and $gU \subset \mathcal{C}$ the image subspace LgU is again Lagrangian and it is the graph of a linear map which we denote by LU i.e., $gLU = LgU$. All Lagrangian subspaces strictly inside the sector are graphs of symmetric positive definite linear maps in $\mathbb{R}^n$. To simplify notation we will write U instead of gU.

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

acts on Lagrangian subspaces by the following Möbius transformation

$$LU = (C + DU)(A + BU)^{-1}.$$

For a monotone L if $U > 0$ then also $LU > 0$.

Proposition 5. *For two positive Lagrangian subspaces defined by* $U : \mathbb{R}^n \to \mathbb{R}^n$ *and*
$W : \mathbb{R}^n \to \mathbb{R}^n$

$$s(U, W) = \max\{\frac{|\ln\lambda|}{2} | \lambda \textit{ is an eigenvalue of} UW^{-1}\} = \sup_{0\neq\xi\in\mathbb{R}^n} \frac{1}{2}|\ln\langle U\xi, \xi\rangle - \ln\langle W\xi, \xi\rangle|.$$

Proof. We have

$$\sinh s(U, W) = \sup_{\substack{0\neq\xi\in\mathbb{R}^n \\ 0\neq\xi'\in\mathbb{R}^n}} \frac{|\langle\xi, W\xi'\rangle - \langle U\xi, \xi'\rangle|}{2\sqrt{\langle\xi, U\xi\rangle}\sqrt{\langle\xi', W\xi'\rangle}}.$$

Putting $\zeta = U^{\frac{1}{2}}\xi$ and $\zeta' = W^{\frac{1}{2}}\xi'$ we get

$$\begin{aligned}
\sinh s(U, W) &= \sup_{\substack{0\neq\zeta\in\mathbb{R}^n \\ 0\neq\zeta'\in\mathbb{R}^n}} \frac{|\langle (W^{\frac{1}{2}}U^{-\frac{1}{2}} - W^{-\frac{1}{2}}U^{\frac{1}{2}})\zeta, \zeta'\rangle|}{2\sqrt{\langle\zeta, \zeta\rangle}\sqrt{\langle\zeta', \zeta'\rangle}} \\
&= \sup_{0\neq\zeta\in\mathbb{R}^n} \frac{\sqrt{\langle (W^{\frac{1}{2}}U^{-\frac{1}{2}} - W^{-\frac{1}{2}}U^{\frac{1}{2}})\zeta, (W^{\frac{1}{2}}U^{-\frac{1}{2}} - W^{-\frac{1}{2}}U^{\frac{1}{2}})\zeta\rangle}}{2\sqrt{\langle\zeta, \zeta\rangle}} \\
&= \sup_{0\neq\zeta\in\mathbb{R}^n} \frac{\sqrt{\langle (U^{-\frac{1}{2}}WU^{-\frac{1}{2}} - 2I + U^{\frac{1}{2}}W^{-1}U^{\frac{1}{2}})\zeta, \zeta\rangle}}{2\sqrt{\langle\zeta, \zeta\rangle}} \\
&= \max\{\frac{\sqrt{\mu}}{2} | \mu \text{ is an eigenvalue of } U^{-\frac{1}{2}}WU^{-\frac{1}{2}} - 2I + U^{\frac{1}{2}}W^{-1}U^{\frac{1}{2}}\} \\
&= \max\{\frac{\sqrt{\lambda - 2 + \lambda^{-1}}}{2} | \lambda \text{ is an eigenvalue of } UW^{-1}\} \\
&= \max\{\frac{|\lambda^{\frac{1}{2}} - \lambda^{-\frac{1}{2}}|}{2} | \lambda \text{ is an eigenvalue of } UW^{-1}\}.
\end{aligned}$$

Which proves the first equality. To prove the last equality we transform further

$$\begin{aligned}
\max\{\frac{|\ln\lambda|}{2} | \lambda \text{ is an eigenvalue of } W^{-\frac{1}{2}}UW^{-\frac{1}{2}}\} &= \sup_{0\neq\zeta\in\mathbb{R}^n} \frac{1}{2}|\ln\frac{\langle W^{-\frac{1}{2}}UW^{-\frac{1}{2}}\zeta, \zeta\rangle}{\langle\zeta, \zeta\rangle}| \\
&= \sup_{0\neq\xi\in\mathbb{R}^n} \frac{1}{2}|\ln\langle U\xi, \xi\rangle - \ln\langle W\xi, \xi\rangle|.
\end{aligned}$$

□

Corollary. *The function* $s(U, W)$ *is a complete metric on the space of positive definite matrices. The topology defined by this metric coincides with the standard topology.*

□

If the least coefficient of expansion of $\mathcal{Q}$ is large then the image of $\mathcal{C}$ under L is narrow. Moreover we have the following formula connecting $\sigma(L)$ and the diameter of $L\mathcal{L}$

Theorem 2. *For a strictly monotone map L the diameter of $L\mathcal{L}$ is equal to the $s-$ distance of LV_1 and LV_2 and*

$$\tanh\left(\frac{s(LV_1, LV_2)}{2}\right) = \frac{1}{\sigma(L)^2}.$$

Proof. $\mathcal{Q}-$ isometries are also isometries when acting on $\mathcal{L}$. Hence by Proposition 3 we can restrict our calculations to

$$L = \begin{pmatrix} I & I \\ T & I+T \end{pmatrix}$$

with diagonal T. We have

$$s(LV_1, LV_2) = \max\{\frac{|\ln\lambda|}{2} | \lambda \text{ is an eigenvalue of } I + T^{-1}\} = \max_{1\le i\le n} \frac{\ln\left(1+t_i^{-1}\right)}{2}$$

where $t_1, \dots, t_n$ are the eigenvalues of T. Now we obtain the formula connecting $s(LV_1, LV_2)$ and $\sigma(L)$ from Proposition 4 by a straightforward calculation.

It remains to estimate above the distance of any two Lagrangian subspaces in $L\mathcal{L}$. So let U and W be two positive definite matrices. Then

$$U_1 = LU = (T + U + TU)(I + U)^{-1} = T + \left(I + U^{-1}\right)^{-1}$$

and the same formula applies to $W_1 = LW$. $U_1W_1^{-1}$ has the same eigenvalues as $W_1^{-\frac{1}{2}}U_1W_1^{-\frac{1}{2}}$ and since

$$I + T > U_1 > T$$

we get

$$W_1^{-\frac{1}{2}}(I+T)W_1^{-\frac{1}{2}} > W_1^{-\frac{1}{2}}U_1W_1^{-\frac{1}{2}} > W_1^{-\frac{1}{2}}TW_1^{-\frac{1}{2}}.$$

The eigenvalues of $W_1^{-\frac{1}{2}}TW_1^{-\frac{1}{2}}$ are the same as the eigenvalues of

$$T^{\frac{1}{2}}W_1^{-1}T^{\frac{1}{2}} = T^{\frac{1}{2}}\left(T + \left(I + W^{-1}\right)^{-1}\right)^{-1}T^{\frac{1}{2}}.$$

But

$$T^{\frac{1}{2}}\left(T + \left(I + W^{-1}\right)^{-1}\right)^{-1}T^{\frac{1}{2}} > T^{\frac{1}{2}}(I+T)^{-1}T^{\frac{1}{2}} = \left(I + T^{-1}\right)^{-1}.$$

Similarily the eigenvalues of $W_1^{-\frac{1}{2}}(I+T)W_1^{-\frac{1}{2}}$ are the same as the eigenvalues of

$$(I+T)^{\frac{1}{2}}W_1^{-1}(I+T)^{\frac{1}{2}} = (I+T)^{\frac{1}{2}}\left(T + \left(I + W^{-1}\right)^{-1}\right)^{-1}(I+T)^{\frac{1}{2}}.$$

Again we can estimate

$$I + T^{-1} = (I+T)^{\frac{1}{2}}T^{-1}(I+T)^{\frac{1}{2}} > (I+T)^{\frac{1}{2}}\left(T + \left(I + W^{-1}\right)^{-1}\right)^{-1}(I+T)^{\frac{1}{2}}.$$

We have established that the eigenvalues of $U_1W_1^{-1}$ are smaller than the eigenvalues of $I+T^{-1}$ and bigger than the eigenvalues of $\left(I+T^{-1}\right)^{-1}$ which gives us the desired estimate. $\square$

§3. The Finsler character of the metric s.

We will be introducing a Finsler metric in the manifold of positive Lagrangian subspaces $\mathcal{L}$ such that all monotone maps act on $\mathcal{L}$ as contractions. In particular since $\mathcal{Q}$-isometries are monotone together with their inverses they act on $\mathcal{L}$ as isometries. Since $\mathcal{Q}$-isometries act transitively on $\mathcal{L}$ such a Finsler metric is uniquely defined by its value on the tangent space at any chosen point in $\mathcal{L}$. We will use the parametrization of $\mathcal{L}$ by positive definite matrices. Since $\mathcal{L}$ is an open subset of the linear space of symmetric matrices all tangent spaces of $\mathcal{L}$ can be identified with the linear space.

We define the Finsler metric in the tangent space at the identity matrix I. For $X \in T_I\mathcal{L}$ its Finsler norm $|X|$ is one half of the usual operator norm, i.e.,

$$|X| = \frac{1}{2}\|X\| = \frac{1}{2}\sup_{0\neq\xi\in\mathbb{R}^n}\frac{\|X\xi\|}{\|\xi\|} = \frac{1}{2}\sup_{0\neq\xi\in\mathbb{R}^n}\frac{|\langle X\xi,\,\xi\rangle|}{\langle\xi,\,\xi\rangle} =$$
$$= \max\{\frac{|\lambda|}{2}|\lambda \text{ is an eigenvalue of } X\}.$$

(Let us note here that the Riemannian metric of the symmetric space is obtained by choosing trX_1X_2 as the scalar product on $T_I\mathcal{L}$.)

Such a definition is correct if and only if the $\mathcal{Q}$-isometries which leave I fixed act as isometries on $T_I\mathcal{L}$ with respect to the chosen norm. The $\mathcal{Q}$-isometries preserving the graph of I have the form

$$\begin{pmatrix} F & 0 \\ 0 & F \end{pmatrix}$$

where F is an orthogonal matrix. The action of such a $\mathcal{Q}$-isometry on $T_I\mathcal{L}$ (the derivative at I of the action on $\mathcal{L}$) is given by the formula

$$T_I\mathcal{L} \ni X \rightarrowtail FXF^{-1} \in T_I\mathcal{L}$$

and indeed

$$|FXF^{-1}| = |X|.$$

Let us summarize the above discussion in the following

Definition 5. *The Finsler metric* $|\cdot|$ *on* $\mathcal{L}$ *is defined by*

$$|X| = \frac{1}{2}\|U^{-\frac{1}{2}}XU^{-\frac{1}{2}}\| = \frac{1}{2}\sup_{0\neq\xi\in\mathbb{R}^n}\frac{|\langle X\xi,\,\xi\rangle|}{\langle U\xi,\,\xi\rangle} =$$
$$= \max\{\frac{|\lambda|}{2}|\lambda \textit{ is an eigenvalue of } XU^{-1}\}$$

for $X \in T_U\mathcal{L}$.

In contrast to the Riemannian metric which is uniquely determined (up to scale) there are infinitely many Finsler metrics of this kind and our choice seems to be rather arbitrary. It is easily justified though by the following

Theorem 3. *The distance function defined by the Finsler metric $|\cdot|$ on $\mathcal{L}$ coincides with the metric s of Definition 4.*

Proof. Since the $\mathcal{Q}$-isometries act transitively on $\mathcal{L}$ and are isometries with respect to both metrics that we are comparing it is sufficient to consider the distances between I and an arbitrary positive definite U. Let $U(t), 0 \leq t \leq 1$, be a path connecting I and U, i.e., $U(0) = I$ and $U(1) = U$. We have

$$s(I, U) = \frac{1}{2} \sup_{0 \neq \xi \in \mathbb{R}^n} \left| \ln \frac{\langle U(1)\xi, \xi \rangle}{\langle \xi, \xi \rangle} \right| = \frac{1}{2} \sup_{0 \neq \xi \in \mathbb{R}^n} \left| \int_0^1 \frac{d}{dt} \ln \langle U(t)\xi, \xi \rangle dt \right| =$$
$$= \frac{1}{2} \sup_{0 \neq \xi \in \mathbb{R}^n} \left| \int_0^1 \frac{\langle U'(t)\xi, \xi \rangle}{\langle U(t)\xi, \xi \rangle} dt \right| \leq \frac{1}{2} \int_0^1 \sup_{0 \neq \xi \in \mathbb{R}^n} \frac{|\langle U'(t)\xi, \xi \rangle|}{\langle U(t)\xi, \xi \rangle} dt$$

which is equal to the length of the path in the Finsler metric. Thus s-distance does not exceed the Finsler distance. We conclude the proof by choosing the path

$$U(t) = e^{tX}$$

where X is defined by $e^X = U$. Its Finsler length is equal to $s(I, U)$. $\square$

The formula (1) implies immediately that with respect to our metric in $\mathcal{L}$ the action of any monotone map is a contraction and morover the contraction can be estimated by the inequality

$$\sinh s(LU, LW) \leq \frac{\sinh s(U, W)}{\sigma(L)^2} \tag{2}$$

which holds for any two positive Lagrangian subspaces U and W.

Let us introduce the coefficient of least contraction of the metric s under the action of a monotone map L

$$\tau(L) = \sup_{\substack{U, W \in \mathcal{L} \\ U \neq W}} \frac{s(LU, LW)}{s(U, W)}.$$

Theorem 4. *The action of a monotone map L on $\mathcal{L}$ equipped with the metric s is a contraction and the coefficient of least contraction $\tau(L)$ is equal to $\frac{1}{\sigma(L)^2}$.*

Proof. The inequality (2) tells us that in the limit of coinciding Lagrangian subspaces the coefficient of contraction is at least $\frac{1}{\sigma(L)^2}$. But our metric is the distance function defined by a Finsler metric so that $\tau(L)$ is equal to the infinitesimal coefficient of least contraction and thus (2) implies immediately that

$$\tau(L) \leq \frac{1}{\sigma(L)^2}.$$

To establish the equality let us assume first that L is strictly monotone. Then by Proposition 4 we can restrict ourselves to

$$L = \begin{pmatrix} I & I \\ T & I+T \end{pmatrix}$$

with diagonal T. The action of such L on positive Lagrangian subspaces is given by the formula

$$\mathcal{L} \ni U \rightarrowtail LU = T + (I + U^{-1})^{-1} \in \mathcal{L}.$$

The derivative of this action at U is

$$T_U\mathcal{L} \ni X \rightarrowtail (I+U)^{-1}X(I+U)^{-1} \in T_LU\mathcal{L}.$$

Hence the infinitesimal coefficient of contraction at $U \in \mathcal{L}$ and $X \in T_U\mathcal{L}$ is equal to

$$\frac{|(I+U)^{-1}X(I+U)^{-1}|}{|X|} = \sup_{0\neq\xi\in\mathbb{R}^n} \frac{\langle X\xi,\, \xi\rangle}{\langle((I+U)T(I+U)+U^2+U)\,\xi,\, \xi\rangle} \left(\sup_{0\neq\xi\in\mathbb{R}^n} \frac{\langle X\xi,\, \xi\rangle}{\langle U\xi,\, \xi\rangle}\right)^{-1}.$$

We know that it does not exceed $\sigma(L)^{-2}$ and we want to show that at least for some U there is an X such that

$$\sup_{0\neq\xi\in\mathbb{R}^n} \frac{\langle X\xi,\, \xi\rangle}{\langle U\xi,\, \xi\rangle} = \sigma(L)^2 \sup_{0\neq\xi\in\mathbb{R}^n} \frac{\langle X\xi,\, \xi\rangle}{\langle((I+U)T(I+U)+U^2+U)\,\xi,\, \xi\rangle}.$$

We will achieve this by choosing $U = uI$ and $X = I$. By straightforward computation we obtain that in this case

$$\frac{|(I+U)^{-1}X(I+U)^{-1}|}{|X|} = \frac{u}{(1+u)(t_1(1+u)+u)}$$

where t_1 is the smallest eigenvalue of T. We find further that the function

$$g(u) = \frac{(1+u)(t_1(1+u)+u)}{u}$$

has a minimum at

$$u = \frac{\sqrt{t_1}}{\sqrt{t_1+1}}$$

which is equal to $(\sqrt{1+t_1}+\sqrt{t_1})^2 = \sigma(L)^2$. This proves the Theorem for strictly monotone maps. To extend it to all monotone maps we proceed in the same way as at the end of Proposition 3. For any $\epsilon > 0$ we choose a strictly monotone map L_ϵ so close to the identity that

$$\frac{1}{\sigma(L_\epsilon L)^2} \geq \frac{1}{\sigma(L)^2} - \epsilon.$$

We have

$$\frac{1}{\sigma(L)^2} - \epsilon \leq \frac{1}{\sigma(L_\epsilon L)^2} = \tau(L_\epsilon L) \leq \tau(L) \leq \frac{1}{\sigma(L)^2}$$

which ends the proof. $\square$

Theorem 4 combined with Theorem 2 yields

$$\tau(L) = \tanh \frac{diam(L\mathcal{L})}{2}$$

for any monotone linear map L. This formula is identical with the formula obtained by Birkhoff [Bir] for the contraction of Hilbert's projective metric under the action of a matrix with all positive entries. (Birkhoff has one fourth of the diameter but it comes from different scaling.)

References

[B 1] P. Bougerol, *Kalman filtering with random coefficients and contractions*, to appear in Siam J. Control & Opt. (1990).

[B 2] P. Bougerol, *Filtre de Kalman Bucy et exposants de Lyapunov*, Lyapunov Exponents, Proceedings, Oberwolfach 1990, L. Arnold, H. Crauel, J.-P. Eckmann (Eds), Lecture Notes in Math. 1486, Springer-Verlag (1991), 112 – 122.

[Bir] G. Birkhoff, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc. **85** (1957), 219 – 227.

[W1] M.P.Wojtkowski, *Invariant families of cones and Lyapunov exponents*, Erg.Th.Dyn.Syst. **5** (1985), 145–161.

[W2] M.P.Wojtkowski, *On uniform contraction generated by positive matrices*, Contemporary Mathematics **50** (1986), 109 – 117.

[W3] M.P.Wojtkowski, *Measure theoretic entropy of the system of hard spheres*, Erg.Th.Dyn.Syst. **8** (1988), 133–153.

[W4] M.P.Wojtkowski, *Systems of classical interacting particles with nonvanishing Lyapunov exponents*, Lyapunov Exponents, Proceedings, Oberwolfach 1990, L. Arnold, H. Crauel, J.-P. Eckmann (Eds), Lecture Notes in Math. 1486, Springer-Verlag (1991), 243 – 262.

[Ves] E. Vesentini, *Invariant metrics on convex cones*, Ann. Sc. Norm. Sup. Pisa ser. 4 **3** (1976), 671 – 696.

Mathematics Department
University of Rome II
Rome Italy
Department of Mathematics
University of Arizona
Tucson AZ 85721