Optimal paths for symmetric actions in the unitary group*

Jorge Antezana, Gabriel Larotonda and Alejandro Varela

Abstract

Given a positive and unitarily invariant Lagrangian $\mathcal{L}$ defined in the algebra of Hermitian matrices, and a fixed interval $[a,b] \subset \mathbb{R}$, we study the action defined in the Lie group of $n \times n$ unitary matrices $\mathcal{U}(n)$ by

$$\mathcal{S}(\alpha) = \int_a^b \mathcal{L}(\dot{\alpha}(t))\, dt\,,$$

where $\alpha : [a,b] \to \mathcal{U}(n)$ is a rectifiable curve. We prove that the one-parameter subgroups of $\mathcal{U}(n)$ are the optimal paths, provided the spectrum of the exponent is bounded by π. Moreover, if $\mathcal{L}$ is strictly convex, we prove that one-parameter subgroups are the unique optimal curves joining given endpoints. Finally, we also study the connection of these results with unitarily invariant metrics in $\mathcal{U}(n)$ as well as angular metrics in the Grassmann manifold. [1]

1 Introduction

The group of $n \times n$ complex unitary matrices $\mathcal{U}(n)$ carries, as any Lie group, a canonical connection without torsion defined on left-invariant vector fields X, Y as $\nabla_X Y = \frac{1}{2}[X,Y]$, whose geodesics are the one-parameter groups $t \mapsto Ue^{tZ}$ (here U is a unitary matrix and Z an anti-Hermitian matrix). We can introduce a Riemannian metric on the unitary group in a standard fashion

$$\langle X, Y \rangle_g = Tr(U^*X(U^*Y)^*) = Tr(XY^*),$$

for U^*X, U^*Y in the Lie algebra of the group, that is, for U^*X, U^*Y anti-Hermitian matrices. It is well-known that the connection just introduced is in fact the Levi-Civita connection of the metric g induced by the trace, and that geodesics are short provided the spectrum of Z is bounded by π (see for instance [3]).
Now consider the bi-invariant Finsler metric given by the spectral norm,

$$\|X\|_U = \|U^*X\| = \|X\|$$

*2010 MSC. Primary 15A18, 51F25; Secondary 47L20, 53C22.

[1]Keywords and phrases: geodesic segment, Lagrangian, optimal path, unitarily invariant norm, unitary group, Grassmann manifold, angular metric.

for any X tangent to a unitary matrix U. Remarkably, if one keeps the connection but changes the metric, the geodesics of the connection are still short for the induced rectifiable distance (which, as in the Riemannian setting, is computed as the infimum of the length of piecewise smooth curves joining given endpoints, and $L(\alpha) = \int_0^1 \|\dot{\alpha}\| dt$). The same result was also proved in [4], using techniques of variational calculus, if the Finsler metrics are given by the p-Schatten norms for $p \geq 2$. This raises a natural question: what do these norms have in common that could imply this phenomenon? A possible answer could be that all these norms are unitarily invariant, thus they induce bi-invariant metrics on the unitary group. One of the main obstacles to deal with general unitarily invariant norms, is that variational arguments become untractable if the norm is not smooth enough.
In this article we prove that this is the right answer, and introduce a new approach that simplifies considerably the technicalities. It is based in a beautiful and deep result due to Thompson on the product of exponential matrices (Theorem 2.1 below).
Our approach also works for more general optimization problems described as follows: fix a bounded interval $[a, b] \subset \mathbb{R}$, and let $\mathcal{S}$ be the action defined on piecewise C^1 curves $\alpha : [a, b] \to \mathcal{U}(n)$ by

$$\mathcal{S}(\alpha) = \int_a^b \mathcal{L}(\dot{\alpha}(t))\, dt,$$

where $\mathcal{L}$ is a Lagrangian defined in the algebra of $n \times n$ matrices, with the following unitary invariance property: for every $n \times n$ matrix A, and every pair of $n \times n$ unitary matrices U and V

$$\mathcal{L}(UAV) = \mathcal{L}(A). \tag{1}$$

As usual, it is asked that the Lagrangian is a convex and positive map, and without loss of generality we will assume that $\mathcal{L}(0) = 0$. A Lagrangian that satisfies these properties will be called *symmetric Lagrangian*. Two classical examples of symmetric Lagrangians are:

- An unitarily invariant norm $\|\cdot\|_\phi$;
- The kinetic energy $E(A) = \|A\|_F^2$, where $\|\cdot\|_F$ denotes the Frobenius norm.

In the first case, we recover the geometric context mentioned above, because the action $\mathcal{S}$ defines the length of α associated to the Finsler structure that considers the norm $\|\cdot\|_\phi$ in each tangent space. Note that in this case, $\mathcal{S}$ does not depend on the parametrization of α. So, there is no significative difference between the problem of finding a curve that minimizes $\mathcal{S}$ among all piecewise C^1 curves or among all piecewise C^1 curves with a given interval of parameters.
However, in the second example, the action associated to the kinetic energy depends on the parametrization. Let $\alpha : [a, b] \to \mathcal{U}(n)$ be a smooth curve. A simple change of variable shows that, if we take the family of curves $\alpha_r : [ra, rb] \to \mathcal{U}(n)$ defined by $\alpha_r(t) = \alpha(t/r)$, then $r \mapsto \mathcal{S}(\alpha_r)$ is a non-increasing function for $r \in (0, +\infty)$. The same phenomenon also

holds for any other convex Lagrangian. This suggests that in order to find a minimum we should fix the length of the interval of parameters. This is also suggested by considering the example of the energy functional, where the parameter t should be interpreted as the time parameter.
As translations of that interval do not change the value of $\mathcal{S}(\alpha)$, without lost of generality we can consider intervals of the form $[0, b]$. So, the optimization problem that we will study is the following:

Problem 1. Given $U, V \in \mathcal{U}(n)$ and $b > 0$, find the piecewise C^1 curves $\gamma : [0, b] \to \mathcal{U}(n)$ such that $\gamma(0) = U$, $\gamma(b) = V$ and γ minimizes the action given by

$$\mathcal{S}(\alpha) = \int_0^b \mathcal{L}(\dot{\alpha}(t))\, dt \tag{2}$$

where $\mathcal{L}$ is a given symmetric Lagrangian.

The second question that arises is whether the minimal paths, when they exist, are unique or not, or if they are unique modulus a reparametrization of the path. Thus we will study the following:

Problem 2. Given $U, V \in \mathcal{U}(n)$, $b > 0$, and a minimizing function $\gamma : [0, b] \to \mathcal{U}(n)$ with $\gamma(0) = U$, $\gamma(b) = V$, is this function the unique minimizer of the Lagrangian for the given endpoints? Is it true that any other minimizing curve with this given endpoints is just a reparametrization of γ?

2 Preliminaries

Throughout this paper $\mathcal{M}_n(\mathbb{C})$ denotes the algebra of complex $n \times n$ matrices, $\mathcal{G}l(n)$ the group of all invertible elements of $\mathcal{M}_n(\mathbb{C})$, $\mathcal{U}(n)$ the group of unitary $n \times n$ matrices, and $\mathcal{H}(n)$ the real subalgebra of Hermitian matrices. If $T \in \mathcal{M}_n(\mathbb{C})$, then $\|T\|$ stands for the usual spectral norm, $|\cdot|$ indicates the modulus of T, i.e. $|T| = \sqrt{T^*T}$, and $\mathrm{tr}(T)$ denotes the trace of T. Given $A \in \mathcal{H}(n)$, $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$ denotes the eigenvalues of A arranged in non-increasing way, and given an arbitrary matrix $T \in \mathcal{M}_n(\mathbb{C})$, $s_1(T) \geq \ldots \geq s_n(T)$ denotes the singular values of T, i.e. the eigenvalues of $|T|$. We will use $\lambda(A)$ (resp. $s(T)$) to denote the vector in $\mathbb{R}^n$ consisting of the eigenvalues of A (resp. the singular values of T). Finally, given $A, B \in \mathcal{H}(n)$, by means of $A \leq B$ we denote that A is less that or equal to B with respect to the Löwner order.

2.1 Product of exponentials

We begin this subsection with the following remarkable result:

Theorem 2.1 (Thompson [17]). *Given $X, Y \in \mathcal{H}(n)$, there exist unitary matrices U and V such that*

$$e^{iX}e^{iY} = e^{i(UXU^*+VYV^*)}\,.$$

We will use the following corollary of Thompson's theorem:

Corollary 2.2. *Let $X, Y, Z \in \mathcal{H}(n)$ be such that $\|Z\| \leq \pi$ and $e^{iX}e^{iY} = e^{iZ}$. Then, there are unitary matrices U and V such that $|Z| \leq |UXU^* + VYV^*|$.*

Proof. By Thompson's Theorem it is enough to prove that, if $X, Y \in \mathcal{H}(n)$, $e^{iX} = e^{iY}$, and $\|X\| \leq \pi$, then $|X| \leq |Y|$. Let $Y = \sum_{n\in\mathbb{N}} \eta_n\ e_n \otimes e_n$ be a spectral decomposition of Y. If $\Lambda = \{n : e^{i\eta_n} = -1\}$, then

$$|X| = \pi P + \sum_{n\notin\Lambda} |\mu_n|\, e_n \otimes e_n\,,$$

where P is the spectral projection of X onto the subspace generated by the eigenvectors associated to $\pm\pi$, and the eigenvalues $\mu_n \in (-\pi, \pi)$ satisfy that $e^{i\mu_n} = e^{i\eta_n}$ for every $n \notin \Lambda$. Clearly $PY = YP$ and $P|X|P \leq P|Y|P$. On the other hand, since $|\mu_n| \leq |\eta_n|$ for every $n \notin \Lambda$, we also obtain that $(1-P)|X|(1-P) \leq (1-P)|Y|(1-P)$. ∎

Another result due to Thompson is the following triangle inequality for the modulus of matrices:

Theorem 2.3 (Thompson [15, 16]). *Given $A, B \in \mathcal{M}_n(\mathbb{C})$, there exist unitaries V and W such that*

$$|X + Y| \leq V|X|V^* + W|Y|W^*.$$

Combining this result with Corollary 2.2 we get:

Proposition 2.4. *Let $m \geq 2$, and consider $X, X_1, \ldots, X_m \in \mathcal{H}(n)$ such that $\|X\| \leq \pi$ and*

$$e^{iX} = e^{iX_1} \cdots e^{iX_m}\,.$$

Then, there exist unitary matrices $U_1, \ldots, U_m$ such that $|X| \leq \sum_{k=1}^{m} U_k|X_k|U_k^*$.

Proof. For $m = 2$ it is a direct consequence of Corollary 2.2 and Theorem 2.3. Suppose that the result is proved for $m = k$. Then, given $X, X_1, \ldots, X_{k+1} \in \mathcal{H}(n)$ such that $\|X\| \leq \pi$, let $Y \in \mathcal{H}(n)$ be such that $\|Y\| \leq \pi$ and

$$e^{iY} = e^{iX_2} \cdots e^{iX_{k+1}}.$$

By the inductive hypothesis, there exist unitary matrices $V_2, \ldots, V_{k+1}$ such that

$$|Y| \leq \sum_{j=2}^{k+1} V_j|X_j|V_j^*\,.$$

On the other hand, since $e^{iX} = e^{iX_1}e^{iY}$, by the case $n = 2$ already proved, there are unitary matrices U_1 and U such that $|X| \leq U_1|X_1|U_1^* + U|Y|U^*$. If we define $U_j = UV_j$ for $j \geq 2$, then we get the desired result. ■

2.2 The Lagrangians

Let us list in the following proposition several properties of the symmetric Lagrangian that will be used in the sequel:

Proposition 2.5. *Let $\mathcal{L} : \mathcal{M}_n(\mathbb{C}) \to [0, \infty)$ be a symmetric Lagrangian, i.e. convex, $\mathcal{L}(0) = 0$, and unitarily invariant in the sense of equation* (1)*. Then*

(P1) $\mathcal{L}$ is continuous,

(P2) $\mathcal{L}(tA) \leq t\mathcal{L}(A)$ for every $t \in [0, 1]$,

(P3) $\mathcal{L}(A) \leq \mathcal{L}(B)$ provided $0 \leq A \leq B$,

(P4) There exists $\phi : \mathbb{R}_+^n \to [0, +\infty)$ such that $\mathcal{L}(A) = \phi(s(A))$. This ϕ is invariant under rearrangement, positive, convex, with $\phi(0) = 0$ and $\phi(x) \leq \phi(y)$ if $x, y \in \mathbb{R}_n^+$ and $x_i \leq y_i$ for $i = 1 \dots n$.

Proof. The first property is clear because every convex function in a finite dimensional vector space is continuous. Also (P2) is a consequence of the convexity and the fact that $\mathcal{L}(0) = 0$. As $\mathcal{L}$ is unitarily invariant, the singular value decomposition implies that $\mathcal{L}(A)$ only depends on the singular values of A. Hence, if $x \in \mathbb{R}_n^+$ and $\mathrm{diag}(x)$ denotes the $n \times n$ diagonal matrix whose diagonal entries correspond to the coordinates of x, we can define $\phi(x) = \mathcal{L}(\mathrm{diag}(x))$; clearly $\phi(0) = 0$, it is non-negative and convex. Convexity implies that if $x, y \in \mathbb{R}_+^n$ and $x_i \leq y_i$ for $i = 1, \dots, n$, then $\phi(x) \leq \phi(y)$. This proves (P4), and (P3) is a direct consequence of it. ■

Remark 2.6. Let $\phi : \mathbb{R}_+^n \to [0, +\infty)$ be a rearrangement invariant, positive and convex function, with $\phi(0) = 0$. Then ϕ gives place to a symmetric Lagrangian $\mathcal{L}_\phi$ via the equation $\mathcal{L}_\phi(A) = \phi(s(A))$. Note that the natural extension of ϕ to $\mathbb{R}^n$ is strongly Schur convex, but not necessarily subadditive.

3 Optimality of one parameter subgroups

A *geodesic segment* is a curve $t \mapsto Ue^{itZ}$ for $Z \in \mathcal{H}(n)$ and $U \in \mathcal{U}(n)$. In this section we prove that the geodesic segments (which are parametrized with constant velocity) are optimal for Problem 1. Moreover, if $\mathcal{L}$ is strictly convex, then we will prove that these geodesic segments are the unique optimal paths.

3.1 Geodesic segments are short

Definition 3.1. *A polygonal path is a broken geodesic, that is, a curve $P : [0, b] \to \mathcal{U}(n)$ such that there is a partition of the interval $[0, b]$ given by the points $0 = t_0 < \ldots < t_k = b$, Herminitian matrices $X_1,\ldots,X_k$ with norm less than or equal to π, and $U \in \mathcal{U}(n)$ so that*

$$P(t) = \begin{cases} Ue^{i\frac{t}{t_1}X_1} & \text{if } t \in [0, t_1] \\ Ue^{iX_1} \cdots e^{iX_{j-1}} e^{i\frac{t-t_{j-1}}{t_j - t_{j-1}}X_j} & \text{if } t \in [t_{j-1}, t_j] \ (j > 1) \end{cases}. \tag{3}$$

Our first step toward the proof of the optimality of the geodesic segments with constant velocity is the following proposition, which proves that segments are better than polygonal paths.

Proposition 3.2. *Let $U \in \mathcal{U}(n)$ and $V = Ue^{iZ}$, with $Z \in \mathcal{H}(n)$ and $\|Z\| \leq \pi$. Let $\gamma : [0, b] \to \mathcal{U}(n)$ be the segment $\gamma(t) = Ue^{it\frac{Z}{b}}$, and $P : [0, b] \to \mathcal{U}(n)$ a polygonal path joining U to V. Then $\mathcal{S}(P) \geq \mathcal{S}(\gamma)$.*

Proof. Let $0 = t_0 < \ldots < t_k = b$, and $X_1,\ldots,X_k \in \mathcal{H}(n)$ with norm less than or equal to π, so that P has the form showed in (3) . Then

$$\begin{aligned} \mathcal{S}(P) &= \sum_{j=1}^{k} \int_{t_{j-1}}^{t_j} \mathcal{L}\big(\dot{P}(t)\big)\, dt = \sum_{j=1}^{k} \int_{t_{j-1}}^{t_j} \mathcal{L}\left(\frac{X_j}{t_j - t_{j-1}}\right) dt \\ &= \sum_{j=1}^{k} (t_j - t_{j-1}) \mathcal{L}\left(\frac{X_j}{t_j - t_{j-1}}\right) \end{aligned} \tag{4}$$

On the other hand, since $e^{iZ} = e^{iX_1} \cdots e^{iX_k}$ and $\|Z\| \leq \pi$, by Proposition 2.4 there exist unitary matrices $U_1, \ldots U_n$ such that

$$|Z| \leq \sum_{k=1}^{n} U_k |X_k| U_k^*. \tag{5}$$

Then, joining (4) and (5), and using the properties of $\mathcal{L}$ we obtain

$$\begin{aligned} \mathcal{S}(P) &= b \sum_{j=1}^{k} \frac{(t_j - t_{j-1})}{b} \mathcal{L}\left(\frac{X_j}{t_j - t_{j-1}}\right) \\ &\geq b\, \mathcal{L}\left(\frac{1}{b} \sum_{j=1}^{k} U_j |X_j| U_j^*\right) \geq b\, \mathcal{L}\Big(\frac{Z}{b}\Big) \\ &= \int_0^b \mathcal{L}\Big(\frac{Z}{b}\Big)\, dt = \mathcal{S}(\gamma). \end{aligned}$$

■

To prove that geodesic segments are optimal paths among all the possible piecewise C^1 curves, we need the following standard approximation result by polygonal paths.

Lemma 3.3. *Let $\alpha : [0, b] \to \mathcal{U}(n)$ be piecewise smooth. Then for any $\epsilon > 0$ there is a polygonal path $P_\epsilon : [0, b] \to \mathcal{U}(n)$ such that for any $t \in [0, b]$,*

$$\|P_\epsilon^*(t)\dot{P}_\epsilon(t) - \alpha^*(t)\dot{\alpha}(t)\| < \epsilon.$$

Proof. We may as well assume that α is smooth in $[0, b]$. Recall that $\alpha, \dot{\alpha}$ are continuous in the uniform norm. Let $\epsilon > 0$, and choose a partition $0 = t_0 < t_1 < \cdots < t_n = b$ of the interval $[0, b]$ such that, for any $k = 0, 1, \cdots, n$,

$$\|\alpha(t) - \alpha(s)\| < 2 \quad \text{and} \quad \|\alpha^*(t)\dot{\alpha}(t) - \alpha^*(s)\dot{\alpha}(s)\| < \frac{\epsilon}{2}$$

if $s, t \in [t_k, t_{k+1}]$. The first condition implies that there exist $Z_k \in \mathcal{H}(n)$ such that $\|Z_k\| < \pi$ and $e^{iZ_k} = \alpha^*(t_k)\alpha(t_{k+1})$. Moreover, if log denotes the principal branch of the logarithm, then

$$Z_k = \log(\alpha^*(t_k)\alpha(t_{k+1})).$$

Now note that, for any fixed $t \in [0, b]$, the map $g : h \mapsto \frac{1}{h}\log(\alpha^*(t)\alpha(t+h))$, is well-defined and analytic, for sufficiently small h. Moreover

$$g(h) \xrightarrow[h\to 0]{} \frac{d}{ds}\log \alpha^*(t)\alpha(t+s)\Big|_{s=0} = \alpha^*(t)\dot{\alpha}(t).$$

Then, taking a refinement of the partition if necessary, we can also assume that

$$\|Z_k - \alpha^*(t_k)\dot{\alpha}(t_k)\| < \frac{\epsilon}{2}$$

for any $k = 0, 1, 2 \cdots, n$. Consider the map $P_\epsilon : [0, b] \to \mathcal{U}(n)$ which is defined as

$$P_\epsilon(t) = \alpha(t_k)e^{\frac{t-t_k}{t_{k+1}-t_k}Z_k} \text{ for } t \in [t_k, t_{k+1}].$$

Then P_ϵ is certainly a polygonal path, and it is straightforward to see that verifies the claim of the lemma. ∎

Theorem 3.4. *Let $U \in \mathcal{U}(n)$ and $V = Ue^{iZ}$, with $Z \in \mathcal{H}(n)$ and $\|Z\| \leq \pi$. Then, the curve $\gamma(t) = ue^{itZ/b}$ is optimal among piecewise smooth curves $\alpha : [0, b] \to \mathcal{U}(n)$ joining U to V, with respect to the action $\mathcal{S}$ defined by a symmetric Lagrangian, and in particular* $\inf S = b\mathcal{L}(Z/b)$.

Proof. Given $\epsilon > 0$, let $\delta > 0$ such that $\|X - Y\| \leq \delta$ implies that $|\mathcal{L}(X) - \mathcal{L}(Y)| < \epsilon/b$ for every X and Y in a ball big enough. Then, let P_δ be a polygonal path in $\mathcal{U}(n)$ as in the previous lemma, joining U to V, such that

$$\|\dot{\alpha} - \dot{P}_\delta\| = \|\alpha^*\dot{\alpha} - P_\delta^*\dot{P}_\delta\| < \delta.$$

Then by Proposition 3.2,

$$\mathcal{S}(\gamma) \leq \mathcal{S}(P_\delta) = \int_0^b \mathcal{L}(\dot{P}(t))\,dt \leq \varepsilon + \int_0^b \mathcal{L}(\dot{\alpha}(t))\,dt < \epsilon + \mathcal{S}(\alpha),$$

Therefore, $\mathcal{S}(\gamma) \leq \mathcal{S}(\alpha)$. ■

Remark 3.5. If $\alpha : [0,b] \to \mathcal{U}(n)$ is just rectifiable (that is, differentiable *p.p.* with $\dot{\alpha}(t)$ bounded), the approximation by a polygonal path can be carried out with no major changes, and the proof of the previous theorem shows that in fact, geodesic segments are optimal among *rectifiable* arcs joining given endpoints.

3.2 Uniqueness of short paths

Concerning uniqueness, it is clear that the convexity condition of $\mathcal{L}$ should be strenghtened. Let us agree to call $\mathcal{L}$ *nondegenerate* if, given $A, B \in \mathcal{H}(n)$, the existence of $\lambda \in (0,1)$ such that the inequality of the convexity condition turns into an equality, implies that there exists $s \geq 0$ such that $A = sB$. In other words, if

$$\mathcal{L}(\lambda A + (1-\lambda)B) = \lambda\mathcal{L}(A) + (1-\lambda)\mathcal{L}(B)$$

for some $\lambda \in (0,1)$, then $A = sB$ for some $s \geq 0$. This is a notion of nondegeneracy outside lines.
The other notion at play here is the strongest notion of *strict convexity* of $\mathcal{L}$, which of course means that if the equality above holds for some $\lambda \in (0,1)$, then $A = B$. A simple example of a strictly convex Lagrangian is the energy functional, given by the square of the Frobenius norm on $\mathcal{H}(n)$.

Remark 3.6. Note that strict convexity implies nondegeneracy, but the notion of nondegeneracy is relevant since no linear space norm can be strictly convex. In fact, it is usual to say that a norm $\|\cdot\|$ on a linear space is strictly convex when the weaker condition (nondegeneracy) stated above holds, which due to the homogeneity of the norm amounts to say that

$$\|A + B\| = \|A\| + \|B\|$$

implies $A = sB$ for some $s \geq 0$, and geometrically, is equivalent to the fact that the unit ball of the normed space has no segments.

We begin with a technical lemma. Recall that if $A \in \mathcal{H}(n)$, then $\lambda_1(A), \ldots, \lambda_n(A)$ denotes the eigenvalues of A arranged in non-increasing way.

Lemma 3.7. *Let $X, Y, Z \in \mathcal{H}(n)$ be such that $e^Z = e^{iX}e^{iY}$ and $\|Z\| < \pi$. If $\lambda_k(X) = r\lambda_k(Z)$ and $\lambda_k(Y) = (1-r)\lambda_k(Z)$ for some $r \in [0,1]$ and every $k \in \{1,\ldots,n\}$, then $X = rZ$ and $Y = (1-r)Z$.*

Proof. It is enough to show that Z shares an orthonormal basis of eigenvalues with X and Y. Let ξ be an unitary eigenvector of Z such that $|Z|\xi = \|Z\|\xi$. Consider the unit sphere $S^{n-1} \subset \mathbb{C}^n$ and the maps $\alpha, \beta : [0,1] \to S^{n-1}$ given by $\alpha(t) = e^{itZ}\xi$,

$$\beta(t) = \begin{cases} e^{2itX}\xi & \text{if } t \in [0, 1/2] \\ e^{iX}e^{2i(t-1/2)Y}\xi & \text{if } t \in [1/2, 1] \end{cases}.$$

In particular, α and β have the same extreme points. A simple computation shows that, with respect to the natural Riemannian structure, $\mathrm{Long}(\alpha) = \mu$ and $\mathrm{Long}(\beta) \leq \mu$. But, since

$$\ddot{\alpha}(t) = e^{itZ}(-Z^2)\xi = -e^{itZ}|Z|^2\xi = -\|Z\|^2 e^{itZ}\xi = -\|Z\|^2\alpha(t)$$

and $\mathrm{Long}(\alpha) = \|Z\| < \pi$, then α is the unique short geodesic of the sphere S^{n-1} joining ξ with $e^{iZ}\xi$. So, $\mathrm{Graph}(\alpha) = \mathrm{Graph}(\beta)$ and ξ is also an eigenvalue of X and Y. Iterating this procedure, we can conclude that X, Y and Z share a common orthonormal basis of eigenvalues. ■

Theorem 3.8. *Assume that $\mathcal{L}$ is strictly convex. Let $X, Y \in \mathcal{H}(n)$ with norm less or equal than π, and $Z \in \mathcal{H}(n)$ such that $\|Z\| < \pi$ and $e^{iZ} = e^{iX}e^{iY}$. Consider the geodesic segment $\gamma : [0, b] \to \mathcal{U}(n)$ defined by $\gamma(t) = e^{itZ/b}$, and the polygonal $P : [0,1] \to \mathcal{U}(n)$defined by*

$$\begin{cases} e^{i\frac{t}{t_0}X} & \text{if } t \in [0, t_0] \\ e^{iX}e^{i\frac{t-t_0}{b-t_0}Y} & \text{if } t \in [t_0, b] \end{cases}.$$

for some $t_0 \in (0, b)$. If $\mathcal{S}(P) = \mathcal{S}(\gamma)$ then $X = \frac{t_0}{b}Z$ and $P = \gamma$.

Proof. By Proposition 2.1, there exist unitary matrices U and V such that

$$e^{iZ} = e^{i(UXU^* + VYV^*)} \quad \text{and} \quad |Z| \leq |UXU^* + VYV^*|,$$

and by the computations made in Proposition 3.2 (Equation (4))

$$\mathcal{S}(P) = t_0\,\mathcal{L}\left(\frac{X}{t_0}\right) + (b - t_0)\,\mathcal{L}\left(\frac{Y}{b - t_0}\right).$$

Then, using the properties of $\mathcal{L}$, the hypothesis $\mathcal{S}(P) = \mathcal{S}(\gamma)$ implies that

$$\begin{aligned} \mathcal{S}(\gamma) = \mathcal{S}(P) &= t_0\,\mathcal{L}\left(\frac{X}{t_0}\right) + (b - t_0)\,\mathcal{L}\left(\frac{Y}{b - t_0}\right) \\ &= b\left(\frac{t_0}{b}\,\mathcal{L}\left(\frac{UXU^*}{t_0}\right) + \frac{b - t_0}{b}\,\mathcal{L}\left(\frac{VYV^*}{b - t_0}\right)\right) \\ &\geq b\mathcal{L}\left(\frac{UXU^* + VYV^*}{b}\right) \geq b\,\mathcal{L}\Big(\frac{Z}{b}\Big) = \mathcal{S}(\gamma). \end{aligned}$$

On one hand, this implies that $Z = UXU^* + VYV^*$. Indeed, if $W = UXU^* + VYV^*$ then $|Z| \leq |W|$. But the above chain of identities implies that $\mathcal{L}(Z) = \mathcal{L}(W)$, and (P2) in Proposition 2.5 implies that $|Z| = |W|$. Hence, $0 \leq |Z| = |W| < \pi$. Since $e^{iZ} = e^{i(UXU^*+VYV^*)}$ we get the desired equality. On the other hand, since $\mathcal{L}$ is strictly convex if $r = t_0/b$ then

$$rZ = UXU^* \quad \text{and} \quad (1-r)Z = VYV^*.$$

Now, by Lemma 3.7 we obtain that $X = UXU^*$ and $Y = VYV^*$ which concludes the proof. ■

Theorem 3.9. *Assume that $\mathcal{L}$ is strictly convex. Let $Z \in \mathcal{H}(n)$ be such that $\|Z\| < \pi$. Then, the geodesic segment $\delta : [0, b] \to \mathcal{U}(n)$ defined by $\gamma(t) = Ue^{itZ/b}$ is the unique piecewise C^1 curve in $\mathcal{U}(n)$ joining U to $V = Ue^{iZ}$, and $\mathcal{S}(\delta) = b\mathcal{L}(Z/b)$.*

Proof. Without lost of generality we can assume that $U = 1$. Suppose that α is any short, piecewise smooth curve joining 1 to e^{iZ}. Let $t_0 \in (0, 1)$ and let $\alpha(t_0) = e^{iX} = e^{iZ}e^{-iY}$, with $\|Y\| \leq \pi$, $\|X\| \leq \pi$. Consider the polygonal $P : [0, b] \to \mathcal{U}(n)$ defined by

$$\begin{cases} e^{i\frac{t}{t_0}X} & \text{if } t \in [0, t_0] \\ e^{iX}e^{i\frac{t-t_0}{b-t_0}Y} & \text{if } t \in [t_0, b] \end{cases}.$$

Then, by Proposition 3.2 and Theorem 3.4 applied to each segment,

$$\mathcal{S}(\gamma) \leq \mathcal{S}(P) \leq \int_0^{t_0} \mathcal{L}(\dot{\alpha})\, dt + \int_{t_0}^{b} \mathcal{L}(\dot{\alpha})\, dt = \mathcal{S}(\alpha) = \mathcal{S}(\gamma),$$

Hence $\mathcal{S}(\gamma) = \mathcal{S}(P)$, and by Theorem 3.8 we get that $X = \frac{t_0}{b}Z$. ■

This settles Problem 2 when the Lagrangian is strictly convex: the geodesic segments are optimal and unique as functions. Regarding the second question of that problem, we have the following result, that settles this poblem when the Lagrangian is nondengenerate (for instance, if $\mathcal{L}$ is a strictly convex norm on a linear space, Remark 3.6): in this case, geodesic segments are optimal and unique modulo a reparametrization of the path, that is, they are unique in a geometrical sense.

Theorem 3.10. *Assume that $\mathcal{L}$ is nondegenerate. Let $Z \in \mathcal{H}(n)$ be such that $\|Z\| < \pi$. Then, if $\alpha : [0, b] \to \mathcal{U}(n)$ is an optimal path of the minimization problem given by $\mathcal{L}$ with given endpoints U, V, α must be a reparametrization of the geodesic segment $\gamma : [0, b] \to \mathcal{U}(n)$ defined by $\gamma(t) = Ue^{itZ/b}$.*

Proof. We assume that $U = 1$ and $V = e^{iZ}$. Let $t_0 \in (0, 1)$ and let $\alpha(t_0) = e^{iX} = e^{iZ}e^{-iY}$, with $\|Y\| \leq \pi$, $\|X\| \leq \pi$. Arguing as in the proof of Theorem 3.8, convexity of $\mathcal{L}$ and

minimality of α imply that $Z = UXU^* + VYV^*$. Now, nondegeneracy of $\mathcal{L}$ implies also that there exists $s \geq 0$ such that

$$\frac{UXU^*}{t_0} = s\frac{VYV^*}{b - t_0}.$$

Now we take $s_0 = \frac{st_0}{b-t_0} \geq 0$ and $r = (1 + s_0)^{-1}$. Note that $r \in [0, 1]$ and also that $rZ = UXU^*$, $(1 - r)Z = VYV^*$. Invoking once again Lemma 3.7, it follows that $X = UXU^*$, $Y = VYV^*$. Thus $\alpha(t_0) = e^{irZ}$ and then α must be a reparametrization of the geodesic segment γ. ■

Regarding uniqueness of paths when $\|U - V\| = 2$ (or equivalently, when $V = Ue^{iZ}$ and $\|Z\| = \pi$), this property is not expected since taking $n = 1$, $U = 1$, $V = -1$ shows that there are two geodesic segments in the circumference ($= \mathcal{U}(1)$) joining U, V, and the situation worsens as n gets bigger.

4 Rectifiable distances in $\mathcal{U}(n)$ and angular metrics in the Grassmann manifold

In this section, we focus in the particular case where $\mathcal{L}$ is a unitarily invariant norm. In that case the action $\mathcal{S}$ defines a length of curves and the length of the optimal path defines a distance in $\mathcal{U}(n)$.

4.1 Unitarily invariant norms and symmetric gauge functions

One of the most relevant properties of the uniform norm of matrices is the following: given two unitary matrices U and V, then $\|UTV\| = \|T\|$. This property is shared by many other norms defined in $\mathcal{M}_n(\mathbb{C})$.

Definition 4.1. *A norm $|||\cdot|||$ defined in $\mathcal{M}_n(\mathbb{C})$ is called unitarily invariant if for every matrix T and every pair of unitary matrices U and V it holds that $|||\,UTV\,||| = |||\,T\,|||$.*

As a consequence of the singular value decomposition, $|||\,T\,||| = |||\,|T|\,|||$, and

$$|||\,T\,||| = \|T\|_\phi = \phi(s(T))\,, \tag{6}$$

where ϕ is a *symmetric gauge function*, that is, a rearrangement invariant norm on $\mathbb{R}^n$, and depends only on the moduli of the coordinates of the vectors. The next theorem [5] will be useful in what follows:

Theorem 4.2. *There is a bijection bewtween symmetric gauge functions ϕ on $\mathbb{R}^n$, and unitarily invariant norms $\|\cdot\|_\phi$ on $\mathcal{M}_n(\mathbb{C})$ given by equation (6) above.*

4.2 Rectifiable metrics in the unitary group

By considering as a Lagrangian a unitarily invariant norm $\|\cdot\|_\phi$, the action S can be interpreted as the length of curves L_ϕ, and the rectifiable distance between $U, V \in \mathcal{U}(n)$ is

$$d_\phi(U,V) = \inf\left\{L_\phi(\gamma) |\, \gamma : [a,b] \to \mathcal{U}(n) \text{ is piecewise smooth and joins } U \text{ to } V \text{ in } \mathcal{U}(n)\right\}.$$

The function d_ϕ is in fact a distance, since $\|U-V\|_\phi \simeq d_\phi(U,V)$ for any $U, V \in \mathcal{U}(n)$. One of the main features of this metric is that it is invariant for the action of the unitary group $\mathcal{U}(n)$, in fact it is a bi-invariant metric

$$d_\phi(UV_1W, UV_2W) = d_\phi(V_1, V_2)$$

for $U, W, V_1, V_2 \in \mathcal{U}(n)$.

4.2.1 Minimality of one-parameter subgroups

As a direct consequence of Theorem 3.4 and Theorem 3.10, we obtain the following result, which generalizes [4, Theorem 3.2] for the p-norms ($p \geq 2$), see also [11].

Theorem 4.3. *Let $U, V \in \mathcal{U}(n)$ and $V = Ue^{iZ}$, with $\|Z\| \leq \pi$, $Z \in \mathcal{H}(n)$. Then, the curve $\delta(t) = Ue^{itZ}$ is shorter than any other piecewise smooth curve γ in $\mathcal{U}(n)$ joining U to V, when we measure them with the norm $\|\cdot\|_\phi$. In particular, $d_\phi(U,V) = \|Z\|_\phi$. If $\|U-V\| < 1$ (equivalently, if $\|Z\| < \pi$), then this δ is the unique short path joining U, V in $\mathcal{U}(n)$ provided the norm is stricly convex.*

Remark 4.4. A question related to the uniqueness of geodesics, is if we can ensure that the points in $\mathcal{U}(n)$ are aligned when the distance is additive. That is, if

$$d_\phi(U,V) = d_\phi(U,W) + d_\phi(W,V).$$

implies that there exists $t_0 \in [0,1]$ and $X_0 \in \mathcal{H}(n)$ with $\|X_0\| \leq \pi$ such that

$$V = Ue^{iX_0}, \qquad \text{while} \quad W = Ue^{it_0X_0}.$$

The previous theorem implies this when $\|U-V\| < 2$. However, the question always has an affirmative answer (provided the norm is strictly convex), with a simpler proof.

Theorem 4.5. *Assume that the norm $\|\cdot\|_\phi$ is strictly convex, and let $U, V, W \in \mathcal{U}(n)$ be such that*

$$d_\phi(U,V) = d_\phi(U,W) + d_\phi(W,V).$$

Then U, V, W are aligned in $\mathcal{U}(n)$.

Proof. We can assume that $U = 1$, $V = e^{iZ}$, $W = e^{iX}$ with X, Z of norm less or equal than π. Let $Y \in \mathcal{H}(n)$ such that $\|Y\| \leq \pi$ and $e^{iZ} = e^{iX}e^{iY}$. Then the hypothesis is that

$$\|Z\|_\phi = \|X\|_\phi + \|Y\|_\phi.$$

Consider the smooth path $\alpha(t) = e^{itX}e^{itY}$. Then α joins the same endpoints that $\delta(t) = e^{itZ}$ in $\mathcal{U}(n)$, thus

$$\|X + Y\|_\phi = L_\phi(\alpha) \geq L_\phi(\delta) = \|Z\|_\phi = \|X\|_\phi + \|Y\|_\phi.$$

Since the norm is strictly convex, there exists $\lambda \geq 0$ such that $Y = \lambda X$. Pick $X_0 = (1+\lambda)X$ and $t_0 = (1 + \lambda)^{-1}$ to finish the proof. ■

4.3 The Grassmannian

The Grassmannian $\mathcal{G}_n$ is the set of subspaces of $\mathbb{C}^n$, which can be identified with the set of orthogonal projections in $\mathcal{M}_n(\mathbb{C})$. If we consider in $\mathcal{M}_n(\mathbb{C})$ the topology defined by any of all the equivalent norms, the Grassmann space endowed with the inherited topology becomes a compact set. However, it is not connected. Indeed, it is enough to consider the trace tr, which is a continuous map defined on the whole space $\mathcal{M}_n(\mathbb{C})$, and restricted to $\mathcal{G}_n$ takes only positive integer values. In particular, this shows that the connected components of $\mathcal{G}_n$ are the subsets $\mathcal{G}_{m,n}$ defined as:

$$\mathcal{G}_{m,n} := \{P \in \mathcal{G}_n : \ \mathrm{tr}(P) = m\}.$$

Each of these components is a submanifold of $\mathcal{M}_n(\mathbb{C})$ [18, p.129], and connected components are given by the unitary orbit of a given projection P such that $tr(P) = m$:

$$\mathcal{G}_{m,n} = \{UPU^* : U \in \mathcal{U}(n)\}.$$

The tangent space at a point $P \in \mathcal{G}_{m,n}$ can be identified with the subspace of P-codiagonal Hermitian matrices, i.e.

$$T_P\mathcal{G}_n = \{X \in \mathcal{H}(n) : \ X = PX + XP\} \ .$$

In particular note that $T_P\mathcal{G}_n$ has a natural complement N_P, which is the space of Hermitian matrices that commute with P, that is, the P-diagonal Hermitian matrices. The decomposition in diagonal and codiagonal matrices defines a normal bundle, and leads to a covariant derivative

$$\nabla_V \Gamma(P) = \Pi_{T_P||N_P} \frac{d}{dt}\Gamma(\alpha(t))\bigg|_{t=0}, \tag{7}$$

where Γ is a vector field along the curve $\alpha : (-\varepsilon, \varepsilon) \to \mathcal{G}_{m,n}$ that satisfies $\alpha(0) = P$ and $\dot{\alpha}(0) = V$. So, we have a notion of parallelism, and the geodesics in this sense are described by the following theorem:

Theorem 4.6 (Porta-Recht [13]). *The unique geodesic at P with direction X is:*

$$\gamma(t) = e^{itX} P e^{-itX}.$$

As the unitary group acts transitively in these components via $U \cdot P = UPU^*$, they are also homogeneous spaces of $\mathcal{U}(n)$. They can be distinguished from other homogeneous submanifolds of $\mathcal{U}(n)$, because the map

$$P \mapsto S_P = 2P - 1$$

embeds them in $\mathcal{U}(n)$, and the map S is two times an isometry. The images S_P are symmetries, i.e. matrices that satisfy $S_P^* = S_P = S_P^{-1}$.

4.3.1 Finsler metrics on the Grassmannian

For a given symmetric norm, the Grassmann space carries the Finsler structure given by

$$\|X\|_P = \|X\|_\phi$$

for $X \in T_P\mathcal{G}_n$, and with this structure, the Grassmann component $\{UPU^* : U \in \mathcal{U}(n)\}$ is isometric (modulo a factor 2) to the orbit of symmetries $\{US_PU^* : U \in \mathcal{U}(n)\}$. In the particular case when $\|\cdot\|_\phi$ is the Frobenius norm, this connection is the Levi-Civita connection of the metric, since the P-diagonal matrices are the orthogonal complement of the P-codiagonal matrices with respect to this Riemannian metric.

A straightworward computation shows that, if $X = XP + PX$, then $e^{iX}S_P = S_P e^{-iX}$. This simple observation enables to use our results in the unitary group, to prove minimality of geodesics in the Grassmann manifold:

Theorem 4.7. *If $P, Q \in \mathcal{G}_{m.n}$ then there exists $X \in T_P\mathcal{G}_n$ such that $Q = e^{iX}Pe^{-iX}$ and $\|X\| \leq \frac{\pi}{2}$, unique when $\|P - Q\| < 1$. The geodesic $\gamma(t) = e^{itX}Pe^{-itX}$ is shorter than any rectifiable path in $\mathcal{G}_n$ joining P, Q and*

$$d_\phi(P, Q) = \|XP - PX\|_\phi = \|X\|_\phi.$$

If the norm is strictly convex and $\|P - Q\| < 1$, the geodesic is the unique short path joining $P, Q \in \mathcal{G}_n$.

Proof. The existence of X follows from Halmos [8] or Davis and Kahan [6]. Since $e^{2iX} = S_Q S_P$, if $\|Q - P\| < 1$ this X is unique. Since

$$S_{\gamma(t)} = 2\gamma(t) - 1 = e^{itX}S_P e^{-itX} = e^{2itX}S_P = S_P e^{-2itX},$$

and S is two times an isometry, the minimality of γ follows from Theorem 4.3, and the same applies to the uniqueness in the strictly convex case. Finally, $L_\phi(\gamma) = \|XP - PX\|_\phi$, and on the other hand, since $PXP = 0$ then

$$|XP - PX|^2 = |XP + PX|^2 = |X|^2,$$

thus $d_\phi(P, Q) = L_\phi(\gamma) = \||XP - PX|\|_\phi = \||X|\|_\phi = \|X\|_\phi$. ∎

Remark 4.8. In the situation of the previous theorem, it is not hard to see that if $k \in \mathbb{Z}$, then $PX^{2k} = X^{2k}P$, $PX^{2k+1} = -PX^{2k+1}$. Then $P|X| = |X|P = |XP|$ and $(1-P)|X| = |X|(1-P) = |PX|$. Moreover

$$Q = P\cos^2 X + (1-P)\sin^2 X - \frac{i}{2}P\sin 2X + \frac{i}{2}(1-P)\sin 2X,$$

and then $|PQ|^2 = PQP = P\cos^2 X$, which leads to $|PQ| = P\cos X = \cos|XP|$, and likewise $|QP| = (1-P)\cos X = \cos|PX|$. Thus if $Y \in T_p\mathcal{G}_n$ is any other matrix as X, it follows that $P\cos X = P\cos Y$ or equivalently,

$$\cos|XP| = |PQ| = \cos|YP|.$$

4.4 The angular metrics

Let $\mathcal{X}$ and $\mathcal{Y}$ be two m-dimensional subspaces of $\mathbb{C}^n$, and let $P_\mathcal{X}$ and $P_\mathcal{Y}$ be the orthogonal projections onto $\mathcal{X}$ and $\mathcal{Y}$ respectively. *The principal angles* between $\mathcal{X}$ and $\mathcal{Y}$ are the angles $\theta_1(\mathcal{X},\mathcal{Y}),\dots,\theta_m(\mathcal{X},\mathcal{Y}) \in [0,\pi/2)$ whose cosines are the m greatest singular values of $P_\mathcal{X}P_\mathcal{Y}$, see [9].
In [10] Li, Qiu, and Zhang used the principal angles to define metrics in the components of $\mathcal{G}_{m,n}$. Given a symmetric norm $\|\cdot\|_\phi$, they define for $P,Q \in \mathcal{G}_{m,n}$ the following distance:

$$\rho_\phi(P,Q) = \|\arccos|PQ|\|_\phi.$$

These distances are called *angular metrics*, because if ϕ is the symmetric gauge function associated to $\|\cdot\|_\phi$ then

$$\rho_\phi(P,Q) = \phi(\theta_1(\mathcal{X},\mathcal{Y}),\dots,\theta_m(\mathcal{X},\mathcal{Y}),0,\dots,0).$$

where $\mathcal{X} = R(P)$ and $\mathcal{Y} = R(Q)$. The definition of these metrics was motivated not only by pure mathematics but also by engineering applications. For example, in robust control, a linear time-invariant system can be described by a subspace valued frequency function, and the description of an uncertain system needs a suitable distance measure between subspaces. The reader is referred to [10], where other motivations and applications of these metrics are described.

A legitimate question at this point, is if these distances are related to an infinitesimal structure on the manifold $\mathcal{G}_n$, that is, if the angular distance among $P,Q \in \mathcal{G}_{m,n}$ can be computed as the infima of the lengths of the rectifiable arcs joining P,Q. Note that, by Remark 4.8, if X is as in Theorem 4.7, then the angular distance among P,Q can be computed as

$$\rho_\phi(P,Q) = \|\arccos|PQ|\|_\phi = \|XP\|_\phi$$

and this computation does not depend on the particular X. Then, one can be tempted to endow the Grassmannian with the Finsler metric (i.e. tangent norm) given by $\|X\|_P =$

$\|XP\|_\phi$ for $X \in T_P\mathcal{G}_n$. The problem with this definition is that it is not clear how to extended it to the whole $\mathcal{M}_n(\mathbb{C})$ in order to obtain an unitarily invariant norm there.

To this end, it suffices to consider the case $m \leq n/2$. Let ϕ be the symmetric gauge function associated to $\|\cdot\|_\phi$ (see Theorem 4.2), and define $\|\cdot\|_\psi$ in the following way:

$$\|A\|_\psi = \phi\big(1/2(s_1(A) + s_2(A), \ldots, s_{2m-1}(A) + s_{2m}(A), 0, \ldots, 0)\big), \tag{8}$$

where $s_1(A)$,...,$s_n(A)$ denotes the singular values of A counted with multiplicity and ordered in non-increasing way[2]. Straightforward computations show that $\|\cdot\|_\psi$ is a symmetric norm, and also that, for any $Q \in \mathcal{G}_{m,n}$ and $Z \in T_Q\mathcal{G}_n$ it holds

$$\|QZ\|_\phi = \|Z\|_\psi.$$

The following theorem gives the link between the rectifiable distances and the angular metrics:

Theorem 4.9 (Davis-Kahan [6])**.** *Let $P, Q \in \mathcal{G}_{m,n}$, and denote $\mathcal{X} = R(P)$ and $\mathcal{Y} = R(Q)$. Then, if $X \in \mathcal{H}(n)$ is P-codiagonal with $\|X\| \leq \pi/2$ and $Q = e^{iX}Pe^{-iX}$, its spectrum counted with multiplicity is*

$$\big(\pm\theta_1(\mathcal{X},\mathcal{Y}), \ldots, \pm\theta_m(\mathcal{X},\mathcal{Y}), 0 \ldots, 0\big).$$

Consider the rectifiable distance d_ψ associated to the norm given in (8), and take P, Q, X as in Theorem 4.7. Then

$$\begin{aligned} d_\psi(P,Q) &= \|X\|_\psi = \phi\big(1/2(s_1(X) + s_2(X), \ldots, s_{2m-1}(X) + s_{2m}(X), 0, \ldots, 0)\big) \\ &= \phi\big(\theta_1(\mathcal{X},\mathcal{Y}), \ldots, \theta_m(\mathcal{X},\mathcal{Y}), 0 \ldots, 0\big) \\ &= \rho_\phi(P,Q), \end{aligned}$$

by Theorem 4.9, and this establishes the following (obtained by Neretin in [12] with another proof):

Theorem 4.10. *Let $\|\cdot\|_\phi$ be a symmetric norm, and ρ_ϕ its corresponding angular metric in $\mathcal{G}_{m,n}$. Then, there exists an induced symmetric norm $\|\cdot\|_\psi$ such that the corresponding rectifiable distance d_ψ coincides with ρ_ϕ.*

Remark 4.11. In [10, Section 4], the authors prove that when the norm $\|\cdot\|_\phi$ is strictly convex, if the distance among $P, Q, R \in \mathcal{G}_{m,n}$ is additive, then there exists a direct rotation from $\mathcal{X}$ to $\mathcal{Z}$ through $\mathcal{Y}$, where $\mathcal{X} = R(P), \mathcal{Y} = R(Q)$ and $\mathcal{Z} = R(R)$. This last assertion is equivalent to the notion of being aligned as introduced in Remark 4.4. Thus the proof of this fact follows immediatly from Theorem 4.5.

[2]The arithmetic mean can be replaced by any positive mean.

A Appendix: compact operators

The results of the previous sections can be extended to the infinite dimensional setting as follows. Let $\mathcal{H}$ be a complex separable Hilbert space, $\mathcal{B}(\mathcal{H})$ the algebra of bounded operators with the supremum norm, $\mathcal{K}(\mathcal{H})$ the algebra of compact operators, $\mathcal{U}(\mathcal{H})$ the group of unitary operators. Let $\|\cdot\|_\phi : \mathcal{B}(\mathcal{H}) \to \mathbb{R} \cup \{\infty\}$ be a *symmetric norm*, that is a norm such that

$$\|AXB\|_\phi \leq \|A\|\|X\|_\phi\|B\| \tag{9}$$

for $A, X, B \in \mathcal{B}(\mathcal{H})$ (both sides can equal ∞). In particular, it is unitarily invariant, thus it only depends on the singular values of the operator, and as in Theorem 4.2, there is a symmetric gauge function $\phi : \mathbb{R}^\infty \to \mathbb{R}_{\geq 0}$ related to this norm; the relationship is somewhat subtle so we refer the reader to Simon's book [14] for full details on these *symmetrically normed ideals*.

Let $\mathcal{I} \subset \mathcal{K}(\mathcal{H})$ stand for the ideal of operators with finite norm, which will be assumed to be complete with respect to its norm, and let $\mathcal{U}_\phi = \{u \in \mathcal{U}(\mathcal{H}) : u - 1 \in \mathcal{I}\}$. This is a Banach-Lie group, whose Banach-Lie algebra can be readily identified with the anti-Hermitian part of $\mathcal{I}$, that we will denote with $i\mathcal{I}_h$. A straightforward computation using the functional calculus and the fact that $\mathcal{I}$ is an ideal shows that if $\|Z\| \leq \pi$ is self-adjoint and $e^{iZ} = U$, then $Z \in \mathcal{I}$.

A.1 The special unitary groups

The length functional on $\mathcal{U}_\phi$ is defined accordingly as $L_\phi(\alpha) = \int_0^1 \|\dot\alpha\|_\phi$, and the distance d_ϕ is defined as the infima of the lengths of curves in $\mathcal{U}_\phi$ joining given endpoints; in order to prove minimality of geodesic segments, we will need the following extension of Thompson's formula, its proof can be found in [2, Theorem 3.2]:

Theorem A.1. *Given $X, Y \in \mathcal{K}(\mathcal{H})_h$, there is an isometry $w \in B(\mathcal{H})$ ($w^*w = 1$), and unitary operators U and V such that*

$$e^{i\,wXw^*} e^{i\,wYw^*} = e^{i\,U(wXw^*)U^* + i\,V(wYw^*)V^*}\,.$$

Theorem A.2. *Let $U, V \in \mathcal{U}_\phi$, $Z \in \mathcal{I}$ such that $V = Ue^{iZ}$ and $\|Z\| \leq \pi$. Then, the curve $\gamma(t) = Ue^{itZ}$ is minimal among rectifiable curves $\alpha \subset \mathcal{U}_\phi$ joining U, V, with respect to the distance induced by the length L_ϕ, and $d_\phi(U, V) = \|Z\|_\phi$. This curve is unique if the norm is strictly convex and $\|U - V\| < 2$ (equivalently, $\|Z\| < \pi$).*

Proof. If $Z \in \mathcal{I}$ is such that $e^{iZ} = e^{iX}e^{iY}$ and $\|Z\| \leq \pi$ (where we can assume that $X, Y \in \mathcal{I}$), then $e^{iwZw^*} = e^{iwXw^*}e^{iwYw^*}$ for some isometry $w \in \mathcal{B}(\mathcal{H})$ by Theorem A.1. With the same proof as Corollary 2.2, we obtain

$$|wZw^*| \leq |U(wXw^*)U^* + i\,V(wYw^*)V^*|.$$

Due to (9), it follows that

$$\|Z\|_\phi = \|w^*wZw^*w\|_\phi \leq \|wZw^*\|_\phi \leq \|X\|_\phi + \|Y\|_\phi$$

since w is an isometry thus $\|w\| = 1$. Now the rest of the proof of minimality of segments follows as in Section 3. The uniqueness when the norm is strictly convex can be proved invoking Theorem A.1, and arguing as in the proof of Theorem 3.10. ∎

A.2 The restricted Grassmannians

The same considerations hold for the special Grassmannian manifold, whose components can be regarded as unitary orbits of self-adjoint projections $P \in \mathcal{B}(\mathcal{H})$, with the action of these special unitary groups:

$$\mathcal{G}_\phi(P) = \{UPU^* : U \in \mathcal{U}_\phi\}.$$

Since $U - 1 \in \mathcal{I}$, then the orbit is contained in the affine space $P + \mathcal{I}$. Then tangent spaces are identified with

$$T_P\mathcal{G}_\phi(P) = \{X \in \mathcal{I}_h : XP + PX = X\}.$$

A well-known result of Halmos [8] says that if $P, Q \in \mathcal{B}(\mathcal{H})$ are self-adjoint projections whose ranges have the same dimension (including the posiblity of $+\infty$), and the same holds for their kernels, then there exists a P-codiagonal X such that $\|X\| \leq \frac{\pi}{2}$ and $Q = e^{iX}Pe^{-iX}$. Since $\mathcal{G}_\phi \subset P + \mathcal{I}$, it is easy to check that $S_QS_P \in \mathcal{U}_\phi$. Then, $e^{2iX} = S_QS_P$ is also in $\mathcal{U}_\phi$, and it follows that $X \in \mathcal{I}$.

Corollary A.3. *If $P, Q \in \mathcal{G}_\phi(P)$ then there exists $X \in T_P\mathcal{G}_\phi(P)$ such that $Q = e^{iX}Pe^{-iX}$ and $\|X\| \leq \frac{\pi}{2}$, unique when $\|P - Q\| < 1$. The geodesic $\gamma(t) = e^{itX}Pe^{-itx}$ is shorter than any rectifiable path in $\mathcal{G}_\phi(P)$ joining P, Q and $d_\phi(P, Q) = \|XP - PX\|_\phi = \|X\|_\phi$. If the norm is strictly convex and $\|P - Q\| < 1$, the geodesic is the unique short path joining $P, Q \in \mathcal{G}_\phi(P)$.*

Remark A.4. When $\mathcal{I}$ is the ideal of Hilbert-Schmidt operators, the special Grassmannian defined above is known as the *Sato Grassmannian* or the *restricted Grassmannian*. The proof of minimality of one-parameter groups in this Riemann-Hilbert setting was given in [1] with a different technique.

References

[1] E. Andruchow, G. Larotonda, *Hopf-Rinow theorem in the Sato Grassmannian*. J. Funct. Anal. 255 (2008), no. 7, 1692–1712.

[2] J. Antezana, G. Larotonda, A. Varela, *Thompson-type formulae*, preprint arXiv : 1107.0348v1 (2011).

[3] E. Andruchow, *Short geodesics of unitaries in the L^2 metric*. Canad. Math. Bull. 48 (2005), no. 3, 340–354.

[4] E. Andruchow, G. Larotonda, L. Recht, *Finsler geometry and actions of the p-Schatten unitary groups*, Trans. Amer. Math. Soc. 62 (2010), 319-344.

[5] R. Bhatia. Matrix analysis. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.

[6] C. Davis, W. M. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal. 7 (1970) 1–46.

[7] I. C. Gohberg, M. G. Krein. Introduction to the theory of linear nonselfadjoint operators. Translated from the Russian by A. Feinstein. Translations of Mathematical Monographs, Vol. 18 American Mathematical Society, Providence, R.I. 1969.

[8] P. R. Halmos, *Two subspaces*. Trans. Amer. Math. Soc. 144 (1969) 381–389.

[9] C. Jordan, Essai sur la géométrie à n dimensions, Bull. Soc. Math. France, 3 (1875), pp. 103–174.

[10] C.-K. Li, L. Qiu, Y. Zhang. *Unitarily invariant metrics on the Grassmann space*. SIAM J. Matrix Anal. Appl. 27 (2005), no. 2, 507–531 (electronic).

[11] L.E. Mata-Lorenzo, L. Recht. *Convexity properties of* $\mathrm{Tr}[(a^*a)^n]$. Linear Algebra Appl. 315 (2000), no. 1-3, 25–38.

[12] Y. A. Neretin, *On Jordan angles and the triangle inequality in Grassmann manifolds*, Geom. Dedicata 86 (2001), 81–92.

[13] H. Porta, L. Recht, *Minimality of geodesics in Grassmann manifolds*, Proc. Amer. Math. Soc. 100 (1987), no. 3, 464–466.

[14] B. Simon. Trace ideals and their applications. Second edition. Mathematical Surveys and Monographs, 120. American Mathematical Society, Providence, RI, 2005.

[15] R. C. Thompson, *Convex and concave functions of singular values of matrix sums*, Pacific J. Math. 66 (1976), no. 1, 285–290.

[16] R. C. Thompson, *Matrix type metric inequalities*, Linear and Multilinear Algebra 5 (1977/78), no. 4, 303–319.

[17] R.C. Thompson, *Proof of a conjectured exponential formula*. Linear and Multilinear Algebra 19 (1986), no. 2, 187–197.

[18] F. W. Warner, Foundations of differentiable manifolds and Lie groups, Graduate Texts in Mathematics 94, Springer-Verlag, New York-Berlin, 1983.

Jorge Antezana:
Universitat Autónoma de Barcelona.
Departamento de Matemática,
Facultad de Ciencias
Edificio C Bellaterra (08193)
Barcelona, España.
e-mail: jaantezana@mat.uab.cat

Gabriel Larotonda and Alejandro Varela:
Instituto de Ciencias
Universidad Nacional de General Sarmiento.
J. M. Gutiérrez 1150
(B1613GSX) Los Polvorines,
Buenos Aires, Argentina.
e-mails: glaroton@ungs.edu.ar,
avarela@ungs.edu.ar

J. Antezana, G. Larotonda
and A. Varela:
Instituto Argentino de Matemática
"Alberto P. Calderón", CONICET
Saavedra 15, 3er piso
(C1083ACA) Buenos Aires,
Argentina.