International Journal of Mathematics
Vol. 19, No. 10 (2008) 1–24

World Scientific
www.worldscientific.com

GEOMETRY OF UNITARIES IN A FINITE ALGEBRA: VARIATION FORMULAS AND CONVEXITY*

ESTEBAN ANDRUCHOW

Instituto de Ciencias
Universidad Nacional de General Sarmiento
J.M. Gutierrez 1150, (1613) Los Polvorines, Argentina

LÁZARO A. RECHT

Departamento de Matemáticas P. y A.
Universidad Simón Bolívar
Apartado 89000 (1080A) Caracas, Venezuela

Received 6 September 2007

Given a C^*-algebra $\mathcal{A}$ with trace τ, we compute the first and second variation formulas for the p-energy functional F_p of the unitary group $\mathcal{U}_\mathcal{A}$ of $\mathcal{A}$, for $p = 2n$ an even integer, namely:

$$F_p(\gamma) = \int_a^b \tau([\dot\gamma^*\dot\gamma]^n)dt,$$

where $\gamma(t) \in \mathcal{U}_\mathcal{A}$ is a smooth curve for $t \in [a, b]$. As an application of these formulas, we prove that if d_p denotes the geodesic distance of the Finsler metric induced by the p-norm $\|x\|_p = \tau([x^*x]^n)^{1/p}$, $u_0, u_1, u_2 \in \mathcal{U}_\mathcal{A}$ with $\|u_i - u_j\| < \frac{1}{2}\sqrt{2-\sqrt{2}}$ and $\delta(t)$ is a geodesic of $\mathcal{U}_\mathcal{A}$ joining $\delta(0) = u_0$ and $\delta(1) = u_1$, then the mapping

$$f(t) = d_p(u_2, \delta(t))^p, \quad t \in [0, 1]$$

is convex.

Keywords: Unitary element; trace; variation formulas; convexity.

Mathematics Subject Classification 2000: 58B20, 46L05, 46L51, 58B25

1. Introduction

In this paper we study the geometry of the unitary group $\mathcal{U}_\mathcal{A}$ of a C*-algebra $\mathcal{A}$ with a finite trace, endowed with the Finsler metric given by the p-Schatten norms of the trace ($p = 2n$ an even integer). It is a continuation of [2], where it was shown that the unitary group is a complete metric space with the rectifiable metric given by the p-norm, and that any pair of unitaries can be joined by a smooth minimal

*Geometry of unitaries: variation and convexity.

geodesic (if the unitaries lie at norm distance less than 2, the minimal geodesic is unique). We remark that these results do not follow from general facts concerning Riemann or Finsler metrics, because the tangent spaces are not complete for the p-norms, neither is the unitary group a Banach–Lie group with the local structure given by these norms. For background material on C^*-algebras and the differential geometry of infinite-dimensional Finsler manifolds, we refer the reader to the book by Upmeier [11].

In this paper we consider the properties of the (rectifiable) distance function d_p:

$$f(s) = d_p(u_2, \gamma(s))^p,$$

where $u_2 \in \mathcal{U}_\mathcal{A}$ is fixed and γ is a geodesic, namely, a curve of the form $\gamma(s) = u_0 e^{sx}$, with u_0 unitary and x skew-hermitian. To pursue this study, we compute the first and second variational formulas, which we claim are of interest in their own. Using these formulas, we show that the function f is convex provided that u and the endpoints of γ lie at norm distance less than $\frac{1}{2}\sqrt{2-\sqrt{2}}$. Also we give estimates for the Taylor coefficients of f.

One motivation for this study is the problem of the existence of curves of minimal length in homogeneous spaces of the unitary group of a C*-algebra. Namely, in spaces $\mathcal{U}_\mathcal{A}/\mathcal{U}_\mathcal{B}$, where $\mathcal{B} \subset \mathcal{A}$ is a unital sub-C*-algebra. These quotients have natural Finsler metrics introduced in [5], induced by the (quotient) norm. In [6], existence of minimal curves was proved under restrictive conditions on the subalgebra $\mathcal{B}$, which allowed compactness arguments. In order to obtain more general results on minima of the length functional, in the absence of compactness, convexity arguments are useful. More precisely, our main result here shows that the p-power of the distance function d_p from a fixed element u_2 in $\mathcal{U}_\mathcal{A}$ to the left coset $u_1 \cdot \mathcal{U}_\mathcal{B}$ of another near element u_2 in $\mathcal{U}_\mathcal{A}$, is a convex function. Approximation arguments (as $p \to \infty$) should provide insight into the problem of existence of minimal cure.

Denote by $\mathcal{A}_h$ and $\mathcal{A}_{ah}$ the sets of hermitian and skew-hermitian elements of $\mathcal{A}$. The group $\mathcal{U}_\mathcal{A}$ is a Banach–Lie group in the operator norm topology. The Banach–Lie algebra is the space $\mathcal{A}_{ah}$, and the tangent space at an element $u \in \mathcal{U}_\mathcal{A}$ is

$$(T\mathcal{U}_\mathcal{A})_u = u\mathcal{A}_{ah} = \mathcal{A}_{ah}u.$$

If one endows $\mathcal{U}_\mathcal{A}$ with the Finsler metric which consists of the usual norm at each tangent space, it is known that the geodesics (=short curves) are curves of the form $\delta(t) = ue^{tz}$, for $z \in \mathcal{A}_{ah}$ with $\|z\| \leq \pi$. Moreover, any pair of unitaries can be joined by a short curve of this form, which is unique if $\|z\| < \pi$.

We make the assumption that $\mathcal{A}$ has a faithful trace τ. The group $\mathcal{U}_\mathcal{A}$ is a complete topological group when regarded with the 2-metric induced by τ: $\|x\|_2 = \tau(x^*x)^{1/2}$. It is not a Lie group in this topology. Nevertheless one can introduce a non-complete Riemannian metric, endowing the tangent space with this quadratic norm. This type of manifolds (with non complete Riemannian metrics) are called *weak* Riemannian manifolds in the literature. The Levi–Civita connection of this metric is given by the expression:

$$\frac{DX}{dt} = \dot{X} - \frac{1}{2}\{\mu^*\dot{\mu}X + X\mu^*\dot{\mu}\}, \tag{1.1}$$

where μ is a smooth curve and X is smooth a vector field, tangent along μ, i.e. $X_{\mu(t)} \in \mu(t)\mathcal{A}_{ah}$ (here smoothness refers to differentiability in the operator norm topology). Remarkably, the short curves δ described before are geodesics for this connection, that is, they satisfy the Euler equation corresponding to (1.1). In [1], it was proven that they are also short geodesics, when the length of curves is measured using the 2-norm $\| \ \|_2$ at each tangent space, i.e. they are minimal geodesics of the Riemannian metric. Furthermore, in [2], it was proved that they are also minimal for the Finsler metric given by the p-norms $\|x\|_p = \tau((x^*x)^{p/2})^{1/p}$, for $2 \le p \le \infty$.

In Sec. 2, we compute the first and second variation formulas for the p-energy functional F_p, for $p = 2n$ an even integer, namely

$$F_p(\gamma) = \int_0^1 \tau([\dot{\gamma}^*\dot{\gamma}]^n)dt$$

defined on smooth curves $\gamma \in \mathcal{U}_{\mathcal{A}}$ (which, for simplicity, we shall suppose parametrized in the unit interval $[0,1]$).

These formulas will play an important role in the metric study of homogeneous spaces, which we will pursue elsewhere. As an application of these formulas in this direction, we prove our main result in Sec. 4: if d_p denotes the geodesic distance for the Finsler metric induced by the p-norm ($p = 2n$), $\delta(t) \in \mathcal{U}_{\mathcal{A}}$ is a minimizing geodesic joining $\delta(0) = u_0$ and $\delta(1) = u_1$, and u_2 is a fixed element in $\mathcal{U}_{\mathcal{A}}$, then the map

$$f(s) = d_p(u_2, \delta(s))^p, \quad s \in [0,1]$$

is convex, provided that $\|u_i - u_j\| < \frac{1}{2}\sqrt{2-\sqrt{2}}$.

The minimality results for geodesics in the p-norms establish that $\mathcal{U}_{\mathcal{A}}$ with the Finsler p-metric is a geodesic metric space (see [8]). The result above can be formulated in the language of geodesic metric spaces: the map

$$F : \{u \in \mathcal{U}_{\mathcal{A}} : \|u - u_2\| < r/2\} \to \mathbb{R}, \quad F(u) = d_p(u, u_2)^p$$

is convex, for $r = \frac{1}{2}\sqrt{2-\sqrt{2}}$.

In Sec. 5, we estimate the Taylor coefficients of $f(s)$. Namely, if $n \ge 2$ ($p \ge 4$) and (without loss of generality) $u = 1$ and $\delta(s) = e^v e^{sz}$, we prove that

$$f(s) = |\tau(v^p)| + s^p|\tau(z^p)| + \sum_{k=1}^{\infty} Q_k(v,z)s^k,$$

and there exist constants R and $C = C(\|v\|, \|z\|)$ with $C \to 0$ if $\|v\|, \|z\| \to 0$, such that

$$|Q_k(v,z)| \le C\|z\|_2^{1-1/(n-1)} f''(0)^{1/(p-2)},$$

for $\|v\|, \|z\| < R$. In particular, if $f''(0) = 0$, then f reduces to $f(s) = |\tau(v^p)| + s^p|\tau(z^p)|$.

We shall adopt the following conventions. Capital letters X, Y, Z will denote fields of operators. If X is a tangent field, $X_u \in u\mathcal{A}_{ah}$, we will denote with the lower case type x the field left translation of the field to the origin: $x_u = u^* X_u \in \mathcal{A}_{ah}$, and will omit the subindex u when possible. For example, if one performs this translation for the covariant derivative formula (1.1), one obtains, after routine computations:

$$\mu^* \frac{DX}{dt} = \dot{x} - \frac{1}{2}[x, \mu^* \dot{\mu}], \tag{1.2}$$

where $[,]$ is the usual commutator of operators.

Note that an element z of $\mathcal{A}$ which is tangent at $u \in \mathcal{U}_{\mathcal{A}}$ verifies $z^* u + u^* z = 0$. Therefore $z^* = -u^* z u^*$ and $z^* z = -(u^*) z^2$. It follows that we can replace the formula of F_p by the following expression:

$$F_p(\gamma) = (-1)^n \int_0^1 \tau([\gamma^* \dot{\gamma}]^p) dt. \tag{1.3}$$

Note that if γ is a geodesic of $\mathcal{U}_{\mathcal{A}}$, then $v = \gamma^*(t)\dot{\gamma}(t) = (ue^{tz})^*(\dot{ue^{tz}}) = z$ is constant.

Let us also recall the Jacobi operator. For a field X tangent along a curve μ,

$$J(X) = \frac{D^2 X}{dt^2} + R(X, V)V.$$

In our case, following the convention concerning the left translation of the fields,

$$\mu^* J(X) = \mu^* \frac{D^2 X}{dt^2} - \frac{1}{4}[[x, v], v]. \tag{1.4}$$

2. First and Second Variation Formulas for F_p

Let $\gamma_s(t)$, $t \in [0, 1]$, $s \in (-r, r)$ be a smooth variation of the curve γ, i.e.

(1) $\gamma_s(t) \in \mathcal{U}_{\mathcal{A}}$, for all s, t.
(2) The map $(s, t) \mapsto \gamma_s(t)$ is smooth (in the $\| \ \|$ topology of $\mathcal{U}_{\mathcal{A}}$).
(3) $\gamma_0(t) = \gamma(t)$.

Our first task will be to find a formula for

$$\frac{d}{ds} F_p(\gamma_s)|_{s=0}.$$

As in classical differential geometry, we shall call the expression obtained the first variation formula. We adopt the following conventions: $\cdot$ will denote the derivative with respect to the parameter t, and $'$ the derivative with respect to s. Let

$$V_s = \dot{\gamma}_s \quad \text{and} \quad W_s = \gamma_s',$$

and as remarked before, with lower case types the left translations

$$v_s = \gamma_s^* V_s \quad \text{and} \quad w_s = \gamma_s^* W_s.$$

Note that $V_s, W_s \in (T\mathcal{U}_{\mathcal{A}})_{\gamma_s}$ whereas $v_s, w_s \in \mathcal{A}_{ah}$. Then

$$\frac{d}{ds}F_p(\gamma_s) = (-1)^n \frac{d}{ds}\int_0^1 \tau(v_s(t)^p)dt = (-1)^n \int_0^1 \tau\left(\frac{d}{ds}v_s(t)^p\right)dt.$$

By the properties of the trace τ, $\tau(\frac{d}{ds}v_s(t)^p) = p\tau(v_s^{p-1}v_s')$. Note also that

$$v_s' = \frac{d}{ds}\gamma_s^*\dot{\gamma}_s = W_s^* V_s + \gamma_s^* \frac{d}{ds}\dot{\gamma}_s.$$

Therefore

$$\frac{d}{ds}F_p(\gamma_s) = (-1)^n p \int_0^1 \tau(W_s^* V_s v_s^{p-1})dt + (-1)^n p \int_0^1 \tau\left(v_s^{p-1}\gamma_s^* \frac{d}{ds}\dot{\gamma}_s\right)dt. \tag{2.1}$$

Let us first consider the second integral in (2.1):

$$\begin{aligned} (-1)^n p \int_0^1 \tau\left(v_s^{p-1}\gamma_s^* \frac{d}{ds}\dot{\gamma}_s\right) dt &= (-1)^n p \int_0^1 \tau\left(v_s^{p-1}\gamma_s^* \frac{d}{dt}W_s\right)dt \\ &= (-1)^n p \int_0^1 \tau\left(\frac{d}{ds}[v_s^{p-1}\gamma_s^* W_s]\right)dt \\ &\quad - (-1)^n p \int_0^1 \tau\left(\frac{d}{ds}[v_s^{p-1}\gamma_s^*]W_s\right)dt. \end{aligned}$$

Note that $\frac{d}{ds}[v_s^{p-1}\gamma_s^*] = \frac{d}{ds}[v_s^{p-1}]\gamma_s^* + v_s^{p-1}V_s^*$. Therefore

$$\begin{aligned} \frac{(-1)^n}{p}\frac{d}{ds}F_p(\gamma_s) &= \int_0^1 \tau(V_s v_s^{p-1} W_s^*)dt - \int_0^1 \tau(v_s^{p-1}V_s^* W)dt \\ &\quad - \int_0^1 \tau\left(\frac{d}{dt}[v_s^{p-1}]\gamma_s^* W_s\right)dt + \tau(v_s^{p-1}\gamma_s^* W_s)|_{t=0}^{t=1}. \end{aligned}$$

Note that since $p-1$ is odd, v_s^{p-1} is skew-hermitian, and therefore

$$\bar{\tau}(V_s v_s^{p-1} W^*) = -\tau(W_s v_s^{p-1} V_s^*) = -\tau(v_s^{p-1}V_s^* W_s).$$

Then the first two integrals above equal

$$\int_0^1 2Re\ \tau(v_s^{p-1}V_s^* W_s)dt.$$

Since V_s is tangent to $\mathcal{U}_{\mathcal{A}}$ at γ_s, we have that $V_s^* = -\gamma_s^* V_s \gamma_s^*$, as explained before. Then $v_s^{p-1}V_s^* W_s = (\gamma_s^* V_s)^{p-1}(-\gamma_s^* V_s \gamma_s^*)W_s = -v_s^p \gamma_s^* W_s$, where v_s^p is hermitian and $\gamma_s W_s$ is skew-hermitian (recall that W_s is tangent at γ_s). Therefore the trace of this product (of an hermitian times a skew-hermitian operator) is a pure imaginary

1 number. Then the sum of the first two integrals vanishes. If $w_s = \gamma_s^* W_s$, we arrive at the following expression for the first variation:

Theorem 2.1. *Let* $\gamma_s(t) \in \mathcal{U}_{\mathcal{A}}$ *with* $t \in [0,1]$, $s \in (-r,r)$, $v_s(t) = \gamma_s^*(t)\dot\gamma_s(t)$, $w_s(t) = \gamma_s^*(t)\gamma_s'(t)$. *Then*

$$\frac{(-1)^n}{p}\frac{d}{ds}F_p(\gamma_s) = \tau(v_s^{p-1}w_s)|_{t=0}^{t=1} - \int_0^1 \tau\left(\frac{d}{dt}[v_s^{p-1}]w_s\right)dt.$$

Let us compute now the second variation formula. Here we will suppose that the variation γ depends on two parameters s and $\tilde s$: $\gamma = \gamma_{s,\tilde s}$, with $s, \tilde s \in (-r,r)$. We shall compute

$$\frac{d^2}{dsd\tilde s}F_p(\gamma_{s,\tilde s})|_{s=0,\tilde s=0}$$

in the special case when $\gamma_{s,\tilde s}|_{s=0,\tilde s=0}$ is a geodesic of $\mathcal{U}_{\mathcal{A}}$. For brevity, we shall omit the subindices $s, \tilde s$ of the fields v, V, w, W, etc. Let us compute first the derivative of the second term in (2.1):

$$\frac{d}{d\tilde s}|_{\tilde s=0}\int_0^1 \tau\left(\frac{d}{dt}[v^{p-1}]w\right)dt.$$

Note that

$$\frac{d}{d\tilde s}\frac{d}{dt}v^{p-1} = \frac{d}{dt}\frac{d}{d\tilde s}v^{p-1} = \frac{d}{dt}\sum_{k+l=p-2} v^k\left(\frac{d}{d\tilde s}v\right)v^l.$$

Let us denote $\tilde W = \frac{d}{d\tilde s}\gamma$, which is a tangent field along γ, and accordingly $\tilde w = \gamma^*\tilde W$. In general, if Z is a tangent field along γ, by (1.2),

$$\frac{d}{dt}Z = \gamma^*\frac{DZ}{dt} + \frac{1}{2}[z,v] \quad (z = \gamma^* Z).$$

In particular,

$$\begin{aligned}\frac{d}{d\tilde s}v &= \frac{d}{d\tilde s}\gamma^*\frac{d}{dt}\gamma = \left(\frac{d}{d\tilde s}\gamma^*\right)\frac{d}{dt}\gamma + \gamma^*\frac{d^2}{d\tilde s dt}\gamma \\ &= \tilde W^* V + \gamma^*\frac{d}{dt}\tilde W = (\gamma\tilde w)^* V + \gamma^*\frac{d}{dt}(\gamma\tilde w) = \tilde w^* v + v\tilde w + \frac{d}{dt}\tilde w.\end{aligned}$$

3 Note that $\tilde w^* = -\tilde w$. Using that $\frac{d}{dt}\tilde w = \gamma^*\frac{D\tilde W}{dt} - \frac{1}{2}[v,\tilde w]$ (from 1.2), it follows that

$$\frac{d}{d\tilde s}v = [v,\tilde w] + \frac{d}{dt}\tilde w = \frac{1}{2}[v,\tilde w] + \gamma^*\frac{D\tilde W}{dt}. \tag{2.2}$$

In the computation of $\frac{d}{dt}\sum_{k+l=p-2} v^k(\frac{d}{d\tilde s}v)v^l$, we only need to consider the terms $v^k(\frac{d}{dt}(\frac{d}{d\tilde s}v))v^l$. Indeed, the other terms involve derivatives of v with respect to t, which vanish at $s = 0$, $\tilde s = 0$, because $\gamma_{s=0,\ \tilde s=0}(t)$ is a geodesic, and therefore

$v_{s=0,\tilde{s}=0}(t)$ is constant (recall the observation at the end of the first section). It follows that

$$\frac{d}{d\tilde{s}}\frac{d}{dt}v^{p-1}|_{s=0,\tilde{s}=0} = \sum_{k+l=p-2} v^k\left(\frac{d}{dt}\frac{d}{d\tilde{s}}v\right)v^l.$$

Moreover, by (2.2),

$$\frac{d}{dt}\frac{d}{d\tilde{s}}v = \frac{d}{dt}\left(\frac{1}{2}[v,\tilde{w}] + \gamma^*\frac{D\tilde{W}}{dt}\right).$$

Now v is constant with respect to t at $s = 0, \tilde{s} = 0$. Therefore

$$\frac{1}{2}\frac{d}{dt}[v,\tilde{w}] = \frac{1}{2}\left[v, \frac{d}{dt}\tilde{w}\right] = \frac{1}{2}\left[v, \gamma^*\frac{D\tilde{W}}{dt}\right] - \frac{1}{4}[v,[v,\tilde{w}]].$$

Denote by Z the tangent field $\frac{D\tilde{W}}{dt}$. Then the remainder derivative

$$\frac{d}{dt}\left(\gamma^*\frac{D\tilde{W}}{dt}\right) = \frac{d}{dt}z = \gamma^*\frac{DZ}{dt} - \frac{1}{2}[v,z].$$

It follows that

$$\frac{d}{dt}\frac{d}{d\tilde{s}}v = \gamma^*\frac{D^2\tilde{W}}{dt} - \frac{1}{4}[v,[v,\tilde{w}]].$$

Let us remark the ocurrence of the Jacobi operator (1.4) in this computation.

Therefore we arrive at the following formula for the first term in the second variation formula:

$$\begin{aligned}
&\left.\frac{d}{d\tilde{s}}\right|_{s+0,\tilde{s}=0}\int_0^1 \tau\left(\left[\frac{d}{dt}v^{p-1}\right]w\right)dt \\
&= \int_0^1 \tau\left(\sum_{k+l=p-2} v^k\left(\gamma^*\frac{D^2\tilde{W}}{dt^2} - \frac{1}{4}[[\tilde{w},v],v]\right)v^l w\right)dt \\
&= \int_0^1 \tau\left(\sum_{k+l=p-2} v^k(\gamma^* J(\tilde{W})v^l w\right)dt.
\end{aligned}$$

Let us proceed now with the computation of the derivative of the first term in (2.1)

$$\left.\frac{d}{d\tilde{s}}\right|_{s=0,\tilde{s}=0}\tau(v^{p-1}w)|_{t=0}^{t=1}.$$

We shall make a further assumption, namely that $\gamma_{s,\tilde{s}}(0) = 1$. The geometric meaning of this assumption is that the variation γ (putting $s = \tilde{s}$) consists of a family of curves joining 1 to the points of $\delta = \gamma_s(1)$, which is a variation of a given geodesic $\gamma_{0,0}(t)$.

This assumption implies that $w_s(0) = 0$. Recall also the expression (2.2) previously obtained for $\frac{d}{d\tilde{s}}v$,

$$\frac{d}{d\tilde{s}}v = \gamma^* \frac{D\tilde{W}}{dt} + \frac{1}{2}[v, \tilde{w}].$$

In the cases of $\frac{d}{d\tilde{s}}w$ (i.e. $w = v$ in (2.2)) gives

$$\frac{d}{d\tilde{s}}w = \gamma^* \frac{D\tilde{W}}{d\tilde{s}} + \frac{1}{2}[w, \tilde{w}].$$

Then

$$\begin{aligned}\left.\frac{d}{d\tilde{s}}\right|_{s=0,\tilde{s}=0} \tau(v^{p-1})|_{t=0}^{t=1} &= \tau\left(\sum_{k+l=p-2} v^k \gamma^* \frac{D\tilde{W}}{dt} v^l w\right)\Bigg|_{t=0}^{t=1} \\ &\quad + \frac{1}{2}\tau\left(\sum_{k+l=p-2} v^k [v, \tilde{w}] v^l w\right)\Bigg|_{t=0}^{t=1} \\ &\quad + \tau\left(v^{p-1}\left(\gamma^* \frac{D\tilde{W}}{d\tilde{s}} + \frac{1}{2}[w, \tilde{w}]\right)\right)\Bigg|_{t=0}^{t=1}.\end{aligned}$$

Note that elementary computations show that $\tau(\sum_{k+l=p-2} v^k [v, \tilde{w}] v^l w) = \tau(v^{p-1}[w, \tilde{w}])$.

Let us summarize the computation of the second variation, with the given assumptions, putting $s = \tilde{s}$, which implies $W = \tilde{W}$, and therefore the commutants $[w, \tilde{w}]$ vanish.

Theorem 2.2. *Suppose that $\gamma_s(t)$, $t \in [0, 1]$, $s \in (-r, r)$ is a variation of a geodesic γ_0 of $\mathcal{U}_{\mathcal{A}}$, with $\gamma_s(0) = 1$ for all s. Then*

$$\begin{aligned}\left.\frac{d^2}{ds^2}\right|_{s=0} F_p(\gamma_s) &= \tau\left(\sum_{k+l=p-2} v^k \gamma^* \frac{DW}{dt} v^l w\right)_{t=1} \\ &\quad + \int_0^1 \tau\left(\sum_{k+l=p-2} v^k \left(\gamma^* \frac{D^2 W}{dt^2} - \frac{1}{4}[[w, v], v]\right) v^l w\right) dt \\ &\quad + \tau\left(v^{p-1} \gamma^* \frac{DW}{d\tilde{s}}\right)\Bigg|_{t=1} \\ &= \tau\left(\sum_{k+l=p-2} v^k \gamma^* \frac{DW}{dt} v^l w\right)\Bigg|_{t=1} + \int_0^1 \tau\left(\sum_{k+l=p-2} v^k \gamma^* J(W) v^l w\right) dt \\ &\quad + \tau\left(v^{p-1} \gamma^* \frac{DW}{d\tilde{s}}\right)\Bigg|_{t=1}.\end{aligned}$$

Proof. At the extreme $t = 0$, W and w vanish. $\square$

Corollary 2.1. *If additionally $\gamma_s(t)$ consists of a variation of the geodesic γ_0 by geodesics, and furthermore, at $t = 1$, $\gamma_s(1)$ is a geodesic of $\mathcal{U}_{\mathcal{A}}$, then*

$$\left.\frac{d^2}{ds^2}\right|_{s=0} F_p(\gamma_s) = \tau\left(\sum_{k+l=p-2} v^k\gamma^*\frac{DW}{dt}v^l w\right)\Bigg|_{t=1}.$$

Proof. In this case, W is a Jacobi field (in the parameter t), i.e. $J(W) = 0$, and the fact $\gamma_s(1)$ is a geodesic implies that $\frac{DW}{ds} = 0$. $\square$

Remark 2.1. In Sec. 4, we shall apply this formula above to the case of the variation (of geodesics by geodesics) given as follows. Fix $v, z \in \mathcal{A}_{ah}$ such that the unitaries 1, e^v and e^ve^z lie within distance less than 2. Let $\gamma_s(t)$ be the unique geodesic (of the parameter $t \in [0,1]$ joining 1 and e^ve^{sz}, for $s \in [0,1]$, or equivalently, $\gamma_s(t) = e^{t\omega_s}$, where $\omega_s = \log(e^ve^{sz})$. One has that for each fixed s_0, γ_{s_0} is a geodesic, and therefore $\gamma_s(t)$ can be regarded as a variation of this geodesic γ_{s_0}. Therefore the above computation of the derivative holds for s_0 as well.

3. The Hessian of the Power p of the p-Norm

Let us denote by $E_p(x) = \tau((x^*x)^n) = \|x\|_p^p$. In what follows, the Hessian form of this map will play a crucial role. We shall restrict to $\mathcal{A}_{ah}$, therefore $E_p(x) = (-1)^n\tau(x^p)$. It is straightforward to verify that $(dE_p)_x(y) = (-1)^np\tau(x^{p-1}y)$. The second differential equals

$$H_x(y,z) := (d^2E_p)_x(y,z) = (-1)^np\tau\left(\sum_{k+l=p-2} x^kyx^lz\right), \quad x, y, z \in \mathcal{A}_{ah}. \tag{3.1}$$

Note that the the second variation formula can be expressed in terms of H. In [9], Mata-Lorenzo and Recht studied the sign of H, and showed that it is positive semidefinite. Let us transcribe Proposition 3.1 from [9], adapted to our context and notation.

Proposition 3.1. *The form $H_x : \mathcal{A}_{ah} \times \mathcal{A}_{ah} \to \mathbb{R}$ is positive semidefinite. Moreover, it can be given the form*

$$H_x(y) = H_x(y,y) = p\||y||x|^{n-1}\|_2^2 + n\sum_{l+m=n-2}\||x|^l(xy+yx)|x|^m\|_2^2.$$

We shall use this formula later.

Let us prove now that the covariant derivative along a geodesic is compatible with the Hessian form.

Proposition 3.2. *Let $\gamma(t) = e^{tv}$ with $v \in \mathcal{A}_{ah}$, and let X and Y be tangent vector fields along γ, with $x = \gamma^* X$ and $y = \gamma^* Y$ as before. Then*

$$\frac{d}{dt} H_v(x, y) = H_v\left(\gamma^* \frac{DX}{dt}, y\right) + H_v\left(x, \gamma^* \frac{DY}{dt}\right).$$

Proof. Recall from (1.2) that $\dot{x} = \gamma^* \frac{DX}{dt} + \frac{1}{2}[x, \gamma^* \dot{\gamma}]$, where $\gamma^* \dot{\gamma} = v$, and analogously for Y. Then

$$\begin{aligned} \frac{(-1)^n}{p} \frac{d}{dt} H_v(x, y) &= \tau\left(\sum_{k+l=p-2} v^k \left(\gamma^* \frac{DX}{dt} + \frac{1}{2}[x, v]) v^l\right) y\right) \\ &\quad + \tau\left(\sum_{k+l=p-2} v^k x v^l \left(\gamma^* \frac{DY}{dt} + \frac{1}{2}[y, v]\right)\right) \\ &= \frac{(-1)^n}{p} \left\{ H_v\left(\gamma^* \frac{DX}{dt}, y\right) + H_v\left(x, \frac{DY}{dt}\right)\right. \\ &\quad \left. + \frac{1}{2}\tau\left(\sum_{k+l=p-2} v^k [x, v] v^l y + v^k x v^l [y, v]\right)\right\}. \end{aligned}$$

Note that $\sum_{k+l=p-2} v^k [x, v] v^l = x v^{p-1} - v^{p-1} x$, and analogously for y. Then the last term above equals

$$\frac{1}{2}(\tau(x v^{p-1} y - v^{p-1} x y + y v^{p-1} x - v^{p-1} y x)) = 0,$$

which finishes the proof. $\square$

Remark 3.1. In particular, if $X = Y$, we may combine this result with the formula obtained in Corollary 2.1 (and Remark 2.4) to obtain, for a Jacobi field X along a geodesic γ satisfying the hypothesis of Corollary 2.1, the following expression

$$\frac{d^2}{ds^2} F_p(\gamma_s)|_{s=0} = \frac{1}{2} \frac{d}{dt} H_v(x, x)|_{t=1}.$$

Let us compute now the second derivative,

$$\frac{d^2}{dt^2} H_v(x),$$

for the case when $X = \gamma x$ is a Jacobi field along the geodesic $\gamma(t) = e^{tv}$. Recall the Jacobi equation from (1.4):

$$\gamma^* \frac{D^2 X}{dt^2} = \frac{1}{4}[v, [v, x]].$$

Then, from (3.2) we have

$$\frac{d^2}{dt^2}H_v(x) = 2\frac{d}{dt}H_v\left(\gamma^*\frac{DX}{dt}, x\right)$$
$$= 2\left\{H_v\left(\gamma^*\frac{D^2X}{dt}, x\right) + H_v\left(\gamma^*\frac{DX}{dt}, \gamma^*\frac{DX}{dt}\right)\right\}.$$

The first term equals

$$H_v\left(\gamma^*\frac{D^2X}{dt}, x\right) = (-1)^n\frac{p}{4}\tau\left(\sum_{k+l=p-2} v^k[v,[v,x]]v^l x\right) = (-1)^n\frac{p}{4}\tau([v^{p-1},[v,x]]x)$$
$$= (-1)^n\frac{p}{2}\tau(v^p x^2 - v^{p-1}xvx).$$

The second term, after straightforward computations (using the formula for the covariant derivative), equals

$$H_v\left(\gamma^*\frac{DX}{dt}, \gamma^*\frac{DX}{dt}\right) = (-1)^n p\tau\left(\sum_{k+l=p-2} v^k\left(\dot{x} + \frac{1}{2}[v,x]\right)v^l\left(\dot{x} + \frac{1}{2}[v,x]\right)\right)$$
$$= (-1)^n p\left\{\tau\left(\sum_{k+l=p-2} v^k\dot{x}v^l\dot{x} + \sum_{k+l=p-2} v^k\dot{x}v^l[v,x]\right)\right.$$
$$\left. - \frac{1}{2}\tau(v^p x^2 - v^{p-1}xvx)\right\}.$$

Therefore we obtain

$$\frac{d^2}{dt^2}H_v(x,x) = (-1)^n 2p\left\{\tau\left(\sum_{k+l=p-2} v^k\dot{x}v^l\dot{x}\right) + \tau\left(\sum_{k+l=p-2} v^k\dot{x}v^l[v,x]\right)\right\}. \quad (3.2)$$

We can further simplify this expression. Let us write the Jacobi equation in terms of x and its deivatives:

Lemma 3.1. *If X is a Jacobi field along $\gamma(t) = e^{tv}$, then $x = \gamma^* X$ satifies*

$$\ddot{x} + [v, \dot{x}] = 0.$$

Proof. Note that $\frac{DX}{dt} = \gamma(\dot{x} - \frac{1}{2}[x,v])$, therefore if $z = \dot{x} - \frac{1}{2}[x,v]$ and $Z = \gamma z$, one has

$$\frac{D^2X}{dt^2} = \frac{DZ}{dt} = \gamma\left(\dot{z} - \frac{1}{2}[z,v]\right) = \gamma\left(\ddot{x} - \frac{1}{2}[\dot{x},v] - \frac{1}{2}\left[\dot{x} - \frac{1}{2}[x,v], v\right]\right)$$
$$= \gamma\left(\ddot{x} - [\dot{x},v] + \frac{1}{4}[[x,v],v]\right).$$

Therefore the equation $\gamma^* \frac{D^2 X}{dt^2} = \frac{1}{4}[[x,v],v]$ translates into

$$\ddot{x} - [\dot{x}, v] + \frac{1}{4}[[x,v],v] = \frac{1}{4}[[x,v],v],$$

or equivalently

$$\ddot{x} - [\dot{x}, v] = 0. \qquad \square$$

Remark 3.2. This version of the Jacobi equation can be integrated. The solution of

$$\ddot{x} + [v, \dot{x}] = 0, \quad \text{with } x(0) = 0 \text{ and } \dot{x}(0) = \xi_0$$

is given by

$$\dot{x}(t) = e^{-tv}\xi_0 e^{tv}$$

and

$$x(t) = \int_0^t e^{-sv}\xi_0 e^{sv} ds.$$

Therefore,

$$[v, x] = \left[v, \int_0^t e^{-sv}\xi_0 e^{sv} ds\right] = \int_0^1 e^{-sv}[v, \xi_0] e^{sv} ds.$$

Therefore we obtain

Proposition 3.3. *Let $X = \gamma x$ be a Jacobi field along $\gamma(t) = e^{tv}$ $(v \in \mathcal{A}_{ah})$. Then*

$$\frac{d^2}{dt^2} H_v(x,x) = (-1)^n 2p \left\{ \tau \left(\sum_{k+l=p-2} v^k \dot{x} v^l \xi_0 \right) \right\}.$$

Proof. In the expression (3.2)

$$\frac{d^2}{dt^2} H_v(x,x) = (-1)^n 2p \left\{ \tau \left(\sum_{k+l=p-2} v^k \dot{x} v^l (\dot{x} + [v,x]) \right) \right\},$$

note that $\frac{d}{dt}(\dot{x} + [v,x]) = \ddot{x} + [v,\dot{x}] = 0$, i.e. $\dot{x} + [v,x]$ is constant. At $t = 0$, this equals ξ_0. Therefore

$$\frac{d^2}{dt^2} H_v(x,x) = (-1)^n 2p \left\{ \tau \left(\sum_{k+l=p-2} v^k \dot{x} v^l \xi_0 \right) \right\}. \qquad \square$$

4. Convexity Properties of the Geodesic Distance

Let us continue with the notations of the preceeding paragraph. The formula at (3.3) can be written

$$\frac{d^2}{dt^2} H_v(x,x) = 2H_v(\dot{x}, \xi_0). \tag{4.1}$$

Let us denote by $\mathcal{S}(r) = \mathcal{S}_{H_v}(r)$ the sphere of radius r of the form H_v:

$$\mathcal{S}(r) = \{a \in \mathcal{A}_{ah} : H_v(a) = H_v(a,a) = r^2\}.$$

In fact, since H_v may be degenerate, $\mathcal{S}$ perhaps should be called more properly a cylinder. Note that the derivative (of the left translation) of the Jacobi field x can be regarded as a displacement in the sphere $\mathcal{S}(H_v(\xi_0)^{1/2})$, starting at the point ξ_0 at $t=0$. Indeed, recall that $\dot{x}(0) = \xi_0$, and

$$\dot{x}(t) = e^{-tv}\xi_0 e^{tv},$$

so that

$$H_v(\dot{x}) = H_v(e^{-tv}\xi_0 e^{tv}) = H_v(\xi_0),$$

because $\tau(v^k e^{-tv}\xi_0 e^{tv} v^l e^{-tv}\xi_0 e^{tv}) = \tau(e^{-tv} v^k \xi_0 v^l \xi_0 e^{tv}) = \tau(v^k\xi_0 v^l \xi_0)$.

We wish to prove that $\frac{d^2}{dt^2}H_v(x,x) \geq 0$ for all t up to a critical value, wich we also wish to estimate. If $H_v(\xi_0) = 0$, the Cauchy–Schwarz inequality applied to the non negative form H_v implies that $H_v(\dot{x},\xi_0) \equiv 0$ for all t, and therefore, $\frac{d^2}{dt^2}H_v(x,x) \equiv 0$. Thus we need only to consider the case $H_v(\xi_0) > 0$.

In the next propostion, we shall use the following elementary geometric fact. Let Q be a postive semidefinite quadratic form on a vector space $\mathcal{X}$, and denote by $\mathcal{S}_R^Q$ the sphere of radius R of this form, $\mathcal{S}_R^Q = \{x \in \mathcal{X} : Q(x) = R^2\}$. It is in fact a cylinder, if $Q(z) = 0$ and $x \in \mathcal{S}_R^Q$, then the line $\{x + tz : t \in \mathbb{R}\} \subset \mathcal{S}_R^Q$. Suppose that two elements $x_0, x_1 \in \mathcal{S}_R^Q$ satisfy $0 < Q(x_0,x_1) < R^2$. Then the geodesic distance $d_Q(x_0,x_1)$ between these points, measured with the form Q, is given by

$$d_Q(x_0,x_1) = R\arccos(Q(x_0,x_1)/R^2).$$

Proposition 4.1. *With the current notations,*

$$\frac{d^2}{dt^2}H_v(x,x) \geq 0$$

for all

$$0 \leq t \leq \frac{\pi}{2}\frac{H_v(\xi_0)^{1/2}}{H_v([\xi_0,v])^{1/2}},$$

if $H_v([\xi_0,v])^{1/2} \neq 0$. *If* $H_v([\xi_0,v])^{1/2} = 0$, *then* $\frac{d^2}{dt^2}H_v(x,x)$ *is a positive constant.*

Proof. By the equality at the beginning of this section, it suffices to examine the sign of $H_v(\dot{x},\xi_0)$. Suppose first that $H_v([v,\xi_0]) = 0$. By the aforementioned reason, it follows that

$$0 = H_v(e^{-tv}[v,\xi_0]e^{tv}) = H_v([v, e^{-tv}\xi_0 e^{tv}]) = H_v([v,\dot{x}]).$$

Then, since H_v is non negative,

$$\frac{d}{dt}H_v(\dot{x},\xi_0) = H_v(\ddot{x},\xi_0) = H_v(-[v,\dot{x}],\xi_0) = 0.$$

Thus $H_v(\dot{x},\xi_0)$ is constant, and its value at $t=0$ is $H_v(\xi_0) > 0$.

Suppose now $H_v([v,\xi_0]) > 0$. If we regard $\dot{x}$ as displacement in $\mathcal{S}(H_v(\xi_0))^{1/2}$, the quantity $H_v(\dot{x},v)$ which is positive at $t=0$, where it takes the value $H_v(\xi_0) > 0$. It will remain positive as long as $\dot{x}$ reaches the H_v-orthogonal of ξ_0 in the sphere $\mathcal{S}(H_v(\xi_0)^{1/2})$. To reach this point, $\dot{x}$ should travel at least the distance $\frac{\pi}{2}$ times the radius $H_v(\xi_0)^{1/2}$ of the sphere (the length measured with the form H_v). The arc length of $\dot{x}$ up to time t is measured by

$$\int_0^t H_v(\ddot{x})^{1/2}dl = \int_0^t H_v([\dot{x},v])^{1/2}dl = \int_0^t H_v([\xi_0,v])^{1/2}dl = tH_v([\xi_0,v])^{1/2},$$

because $[\dot{x},v] = [e^{-lv}\xi_0 e^{lv}, v] = [\xi_0, e^{-lv}ve^{lv}] = [\xi_0, v]$. It follows that $H_v(\dot{x},\xi_0) \geq 0$ for all t such that

$$tH_v([\xi_0,v])^{1/2} \leq \frac{\pi}{2}H_v(\xi_0)^{1/2},$$

which finishes the proof. □

Remark 4.1. Note that

$$\frac{H_v([\xi_0,v])^{1/2}}{H_v(\xi_0)^{1/2}} \leq \sup\left\{\frac{H_v([z,v])^{1/2}}{H_v(z)^{1/2}} : z \in \mathcal{A}_{ah}, H_v(z) \neq 0\right\} = \|[\,,v]\|_{H_v},$$

where $\|[\,,v]\|_{H_v}$ denotes the norm of the commutator operator $[\,,v]$ acting in $\mathcal{A}_{ah}$ endowed with the norm induced by the form H_v.

In particular, the result above implies that $\frac{d^2}{dt^2}H_v(x,x) \geq 0$ for all $0 \leq t \leq \|[\,,v]\|_{H_v}^{-1}$ (with the convention that this property holds for all $t \geq 0$ if $\|[\,,v]\|_{H_v} = 0$).

Note that this bound $\|[\,,v]\|_{H_v}^{-1}$ depends solely on v. However, it is difficult to compute. Let us provide a bound for t (ensuring $\frac{d^2}{dt^2}H_v(x,x) \geq 0$) in terms of the norm $\|v\|$.

Lemma 4.1. *Let $a, b \in \mathcal{A}_{ah}$. Then*

$$H_a([b,a]) \leq 4\|a\|^2 H_a(b).$$

Proof. Let us recall (3.1), the alternative expression for the quadratic form H_v from [9]

$$H_x(y) = H_x(y,y) = p\||y||x|^{n-1}\|_2^2 + n\sum_{l+m=n-2}\||x|^l(xy+yx)|x|^m\|_2^2.$$

Consider the first term. In our case $x=a$ and $y=[b,a]$:

$$p\||[b,a]||a|^{n-1}\|_2^2 = p\||[b,a]|a^{n-1}\|_2^2.$$

Note that if $x,y \in \mathcal{A}$, $\||x|y\|_2 = \|xy\|_2$. Indeed, we can imbed $\mathcal{A}$ in a finite von Neumann algebra $\mathcal{M}$, with the trace τ extended to a trace (eventually non normal), and perform in $\mathcal{M}$ the polar decomposition $x = u|x|$, where u can be chosen unitary

1st Reading

September 29, 2008 15:24 WSPC/133-IJM 00510

because $\mathcal{M}$ is finite. Then $\||x|y\|_2 = \|u^*xy\|_2 = \|xy\|_2$. Analogously $\|y|x|\|_2 = \|yx\|_2$. Then

$$p\|[b,a]v^{n-1}\|_2^2 \leq p(\|ba^n\|_2 + \|aba^{n-1}\|_2)^2.$$

It is elementary that if $x, y \in \mathcal{A}$, $\|yx\|_2 \leq \|y\|_2\|x\|$ and $\|xy\|_2 \leq \|x\|\|y\|_2$. Then

$$p\||[b,a]||a|^{n-1}\|_2^2 \leq 4p\|a\|^2\|ba^{n-1}\|_2^2 = 4p\|a\|^2\||b||a|^{n-1}\|_2^2,$$

which is the corresponding first term in the expression of $H_a(b)$, times $\|a\|^2$.

Let us consider now the second term

$$n\sum_{l+m=n-2}\||a|^l(a[a,b]+[a,b]a)|a|^m\|_2^2 = n\sum_{l+m=n-2}\|a^l(a[a,b]+[a,b]a)a^m\|_2^2.$$

Note that

$$\begin{aligned}\|a^l(a[a,b]+[a,b]a)a^k\|_2 &= \|a^l(a(ab+ba)-(ab+ba)a^k\|_2\\ &\leq \|a^{l+1}(ab+ba)a^k\|_2 + \|a^l(ab+ba)a^{k+1}\|_2\\ &\leq \|a\|\|a^l(ab+bs)s^k\|_2 + \|a^l(ab+ba)a^k\|_2\|a\|\\ &= 2\|a\|\|a^l(ab+bs)a^k\|_2.\end{aligned}$$

Therefore

$$n\sum_{l+m=n-2}\||a|^l(a[a,b]+[a,b]a)|a|^m\|_2^2 \leq 4\|a\|^2\sum_{l+m=n-2}\||a|^l(ab+ab)|a|^m\|_2^2.$$

The proof follows. □

We can rephrase the last proposition, in terms of $\|v\|$:

Corollary 4.1. *With the current notations,*

$$\frac{d^2}{dt^2}H_v(x,x) \geq 0$$

for all

$$0 \leq t \leq \frac{\pi}{4\|v\|}.$$

Proof. Use the above inequality with $a = v$ and $b = \xi_0$. □

We can now prove our main result.

Theorem 4.1. *Let u_0, u_1 and u_2 in $\mathcal{U}_{\mathcal{A}}$, such that $\|u_i - u_j\| < \frac{1}{2}\sqrt{2-\sqrt{2}} = r$. Let $\delta(t) = u_1e^{tz}$ be the minimal geodesic joining u_1 and u_2. Then $f(s) = d_p(u_0, \delta(s))^p$ (d_p = geodesic distance induced by the p-norm) is a convex function ($s \in [0,1]$).*

Proof. Since the metric is invariant for left translation, we may suppose $u_0 = 1$ without loss of generality. Note that $\|1 - \delta(s)\| < 2r < 2$ for all $s \in [0, 1]$. Indeed, $\|1 - \delta(s)\| \le \|1 - u_1\| + \|u_1 - \delta(s)\| < r + \|1 - e^{sv}\|$, and

$$\begin{aligned}\|1 - e^{sv}\| = r(1 - e^{sv}) &= \sup\{\sqrt{2 - 2\cos(sx)} : ix \in \mathrm{sp}(v)\}\\ &\le \sup\{\sqrt{2 - 2\cos(x)} : ix \in \mathrm{sp}(v)\}\\ &= r(1 - e^{v}) = \|1 - u_2\| < r.\end{aligned}$$

Therefore, for each $s \in [0, 1]$ there exists a unique elemement $w_s \in \mathcal{A}_{ah}$ such that $e^{w_s} = \delta(s)$. Moreover, by the same computation above, the inequality

$$2r = \sqrt{2 - \sqrt{2}} > \|1 - \delta(s)\| = \|1 - e^{w_s}\|$$

implies that $\|w_s\| \le \frac{\pi}{4}$.

Put $\gamma_s(t) = e^{tw_s}$, $t \in [0, 1]$. Note that w_s is a smooth function of the parameter s. Therefore $\gamma_s(t)$ is a variation of geodesics by geodesics, more precisely, for each s_0 in $[0, 1]$, $X_{s_0} := \frac{d}{ds}|_{s_0}\gamma_s$ is a Jacobi field along the geodesic γ_{s_0}, vanishing at $t = 0$ and with the property that $\gamma_s(1) = \delta(s)$ is also a geodesic. In other words, for each $s_0 \in [0, 1]$, this field satisfies the hypothesis of Corollary 2.1. Moreover, for each s, γ_s is has minimal length among curves of unitaries joining 1 and $\delta(s)$ [2]. Therefore

$$f(s) = d(1, \delta(s))^p = \mathrm{length}(\gamma_s)^p = F_p(\gamma_s).$$

Then, writing $x_{s_0} = \gamma_{s_0}^* X_{s_0}$, the formula of Remark 3.1 applies and we obtain

$$\ddot{f}(s_0) = \left.\frac{d^2}{ds^2}\right|_{s_0} F_p(\gamma_s) = \frac{1}{2}\frac{d}{dt} H_{w_{s_0}}(x_{s_0})|_{t=1}.$$

Consider $g(t) = H_{w_{s_0}}(x_{s_0})$. Since $\gamma_s(0) = 1$ for all s, x_{s_0} vanishes at $t = 0$, and therefore $g(0) = 0$. Moreover,

$$\ddot{g} = \frac{d^2}{dt^2} H_{w_{s_0}}(x_{s_0}) \ge 0 \quad \text{for all } t \in \left[0, \frac{\pi}{4\|w_{s_0}\|}\right],$$

and since $\|w_{s_0}\| \le \frac{\pi}{4}$, it follows that this interval includes $t = 1$. It follows that $\dot{g} \ge 0$ for such t, and in particular

$$\dot{g}(1) = \ddot{f}(s_0) \ge 0.$$

$\square$

5. The Taylor Expansion of the p-th Power of the Distance Function

In this section, we study the behaviour near the origin of the convex function

$$f(s) = d_p(u_0, \delta(s))^p$$

of the previous section. Recall that $p = 2n$. Our goal is to study the convex behaviour of this function with more detail. Let us establish first a few facts.

Lemma 5.1. *Fix $0 < \alpha < \pi/2$. If $\|v\| \leq \alpha/2$ then*

$$f''(0) \leq H_v(\xi_0) \leq \frac{f''(0)}{\cos(\alpha)}.$$

In particular, $f''(0) = 0$, if and only if $H_v(\xi_0) = 0$.

Proof. As in the previous section, let $x(t) = \int_0^1 e^{-s\xi_0}\xi_0 e^{sv} ds$ be the solution of the Jacobi equation with $x(0) = 0$ and $\dot{x}(0) = \xi_0$. Then,

$$\begin{aligned} f''(0) = H_v(\dot{x}, x)|_{t=1} &= H_v\left(e^{-v}\xi_0 e^v, \int_0^1 e^{-sv}\xi_0 e^{sv} ds\right) \\ &= H_v\left(\xi_0, \int_0^1 e^{(1-s)v}\xi_0 e^{(s-1)v} ds\right). \end{aligned}$$

Changing variables $l = s - 1$, one has

$$f''(0) = H_v\left(\xi_0, \int_0^1 e^{-lv}\xi_0 e^{lv} dl\right) = \int_0^1 H_v(\xi_0, e^{-lv}\xi_0 e^{lv}) dl = \int_0^1 H_v(\xi_0, \dot{x}(l)) dl.$$

By the Cauchy-Schwarz inequality,

$$|H_v(\xi_0, e^{-lv}\xi_0 e^{lv})| \leq H_v(\xi_0)^{1/2} H_v(e^{-lv}\xi_0 e^{lv})^{1/2} = H_v(\xi_0),$$

and thus

$$f''(0) \leq \int_0^1 |H_v(\xi_0, e^{-lv}\xi_0 e^{lv})| dl \leq H_v(\xi_0).$$

This proves the first inequality. The other inequality is trivial if $H_v(\xi_0) = 0$. Thus we may suppose $H_v(\xi_0) > 0$.

As in Proposition 4.1, we can regard $\dot{x}$ as a curve in the sphere of the form H_v with radius $H_v(\xi_0)^{1/2}$, joining ξ_0 and $\dot{x}$. The length of this curve (measured with the form H_v), up to time t is $tH_v([\xi_0, v])^{1/2}$. We may also suppose that $H_v([\xi_0, v]) > 0$. Otherwise, if $H_v([\xi_0, v]) = 0$, reasoning as in Proposition 4.1, $H_v(\xi_0, \dot{x})$ is constant and equal to its value $H_v(\xi_0)$. Thus by the integral above

$$f''(0) = \int_0^1 H_v(\xi_0, \dot{x}(l)) dl = H_v(\xi_0),$$

and the inequality $H_v(\xi_0) \leq \frac{f''(0)}{\cos(\alpha)}$ holds trivially. Suppose therefore that $H_v([\xi_0, v]) > 0$. The geodesic distance between ξ_0 and $\dot{x}$ (in the sphere of the form H_v) is $\arccos(\frac{H_v(\xi_0, \dot{x})}{H_v(\xi_0)})$. Recall from Lemma 4.1 that $H_v([\xi_0, v]) \leq 4\|v\|^2 H_v(\xi_0)$. Therefore, if $t \leq \frac{\alpha}{2\|v\|}$, then

$$t \leq \alpha \frac{H_v(\xi_0)^{1/2}}{H_v([\xi_0, v])^{1/2}}.$$

Equivalently, $tH_v([\xi_0, v])^{1/2} \leq \alpha H_v(\xi_0)^{1/2}$. The left-hand side quantity in this inequality is the length of the path $\dot{x}$ in the sphere of H_v up to time t, thus it majorizes the geodesic distance between its endpoints in this sphere, i.e.

$$H_v(\xi_0)^{1/2}\arccos\left(\frac{H_v(\xi_0,\dot{x})}{H_v(\xi_0)}\right) \leq \alpha H_v(\xi_0)^{1/2},$$

i.e. $H_v(\xi_0,\dot{x}) \geq \cos(\alpha)H_v(\xi_0)$. Recall that this inequality is valid for all $0 \leq t \leq \frac{\alpha}{2\|v\|}$. Therefore, if $\frac{\alpha}{2\|v\|} \geq 1$, the inequality is valid for all $t \in [0,1]$. Using this inequality in the integral expression of $f''(0)$, one obtains

$$f''(0) = \int_0^1 H_v(\xi_0,\dot{x}(l))dl \geq \cos(\alpha)H_v(\xi_0),$$

which completes the proof. $\square$

Consider as before the variation $\gamma_s(t) = e^{tw_s}$, where $w_s(t) = \log(e^v e^{sz})$. Let us compute $\xi_0 = \dot{x}(0)$ in terms of z and v.

Lemma 5.2. *With the above notations,*

$$\xi_0 = z + \frac{1}{2}[v,z] + \frac{1}{12}[v,[v,z]].$$

Proof. Recall that $x(t) = e^{-tv}\frac{d}{ds}e^{tw_s}|_{s=0}$. Therefore

$$\dot{x}(t) = -ve^{-tv}\frac{d}{ds}e^{tw_s}|_{s=0} + e^{tv}\frac{d}{dt}\frac{d}{ds}e^{tw_s}|_{s=0}.$$

Note that $\frac{d}{dt}\frac{d}{ds}e^{tw_s} = \frac{d}{ds}\frac{d}{dt}e^{tw_s} = \frac{d}{ds}(w_s e^{tw_s})$. At $t=0$, one has that $\frac{d}{ds}e^{tw_s}|_{t=0} = 0$. Therefore

$$\xi_0 = \frac{d}{ds}w_s|_{s=0}.$$

According to the classic Baker–Campbell–Hausdorff formula (see for instance [12]),

$$w_s = \log(e^v e^{sz}) = \sum_{n\geq 1} C_n(v,sz),$$

where

$$C_1(v,sz) = v + sz,$$

$$C_2(v,sz) = \frac{1}{2}[v,sz],$$

$$C_3(v,sz) = \frac{1}{12}[v - sz,[v,sz]],$$

and the next terms involve monomials with a factor s^k for $k \geq 2$. The linear coefficient (considering this expansion as a series in the powers of s), is

$$\xi_0 = z + \frac{1}{2}[v,z] + \frac{1}{12}[v,[v,z]].$$

$\square$

For the next result, we make the assumption that any self-adjoint element of $\mathcal{A}$ can be approximated (in norm) by self-adjoint elements with finite spectrum.

Lemma 5.3. *Let $a, b \in \mathcal{A}_h$, then for any $k \geq 1$,*

$$\|ab\|_2 \leq \|b\|_2^{1-1/k} \|a^k b\|_2^{1/k}.$$

Proof. Suppose first that the spectrum of a is finite, $a = \sum_{i=1}^N \lambda_i p_i$, with p_i pairwise orthogonal projections. Then $\|a^k b\|_2^2 = \tau(a^k b^2 a^k) = \sum_{i=1}^N \lambda_i^{2k} R_i$, where $R_i = \tau(p_i b^2 p_i) \geq 0$ (also note that $\sum_{i=1}^N R_i = \|b\|_2^2$). For a general real N-tuple $\alpha = (\alpha_1, \ldots, \alpha_N)$, and $q \geq 2$, define the weighted norm m_q by

$$m_q(\alpha) = \left(\frac{\sum_{i=1}^N |\alpha_i|^q R_i}{\sum_{i=1}^N R_i} \right)^{1/q}.$$

Note that $m_2(\alpha) = \|ab\|_2 / \|b\|_2$. Therefore, by inequality 2.9.1 in [7], $m_2(\lambda) \leq m_{2k}(\lambda)$, that is,

$$\left(\frac{\|ab\|_2}{\|b\|_2} \right)^{2k} \leq \frac{\|a^k b\|_2^2}{\|b\|_2^2},$$

and therefore $\|ab\|_2 \leq \|b\|_2^{1-1/k} \|a^k b\|_2^{1/k}$. For an arbitrary self-adjoint element a, the proof follows by an elementary approximation argument. $\square$

Lemma 5.4. *Suppose that $n \geq 2$ $(p \geq 4)$. Let $v, z \in \mathcal{A}_{ah}$, and consider the variation $\gamma_s(t) = e^{tw_s}$ and the element ξ_0 as above. Then for v of sufficiently small norm, there exists a constant $K = K(\|v\|, \|z\|)$ depending on the* (*operator*) *norms $\|v\|$ and $\|z\|$ such that*

$$\|zv\|_2 \leq K \|z\|_2^{1-1/(n-1)} H_v(\xi_0)^{1/(p-2)}.$$

Proof. As above, $\xi_0 = \xi_0 = z + \frac{1}{2}[v, z] + \frac{1}{12}[v, [v, z]]$. Recall that H_v is a positive semidefinite real bilinear form. Then

$$H_v(\xi_0) = H_v\left(z + \frac{1}{2}[v, z] + \frac{1}{12}[v, [v, z]] \right) \geq H_v(z) + 2H_v\left(z, \frac{1}{2}[v, z] \right)$$
$$+ 2H_v\left(z, \frac{1}{12}[v, [v, z]] \right) + 2H_v\left(\frac{1}{2}[v, z], \frac{1}{12}[v, [v, z]] \right).$$

By the Cauchy–Schwarz inequality for H_v,

$$H_v(\xi_0) \geq H_v(z) - H_v(z)^{1/2} H_v([v, z])^{1/2}$$
$$- \frac{1}{6} H_v(z)^{1/2} H_v([v, [v, z]])^{1/2} - \frac{1}{12} H_v([v, z])^{1/2} H_v([v, [v, z]])^{1/2}.$$

Recall now the inequality in Lemma 4.1:

$$H_a([b, a])^{1/2} \leq 2\|a\| H_a(b)^{1/2}.$$

Then $H_v([v,z])^{1/2} \leq 2\|v\| H_v(z)^{1/2}$ and succesively $H_v([v,[v,z]])^{1/2} \leq 4\|v\|^2 H_v(z)^{1/2}$. Thus

$$H_v(\xi_0) \geq H_v(z)\left(1 - 2\|v\| - \frac{2}{3}\|v\|^2 - \frac{2}{3}\|v\|^3\right). \tag{5.1}$$

One may choose $\|v\|$ small in order to adjust the constant $M_{\|v\|} = 1 - 2\|v\| - \frac{2}{3}\|v\|^2 - \frac{2}{3}\|v\|^3$ to be as close to 1 as wished. Recall now the expression for the quadratic form H in 3.1:

$$H_x(y) = p\||y||x|^{n-1}\|_2^2 + n \sum_{l+m=n-2} \||x|^l(xy+yx)|x|^m\|_2^2.$$

In particular,

$$H_v(z) \geq 2n\||z||v|^{n-1}\|_2^2. \tag{5.2}$$

In a finite algebra, it is apparent that if $z, v \in \mathcal{A}_{ah}$, $\|zv\|_2 = \||z||v|\|_2$. Let us use now Lemma 5.3:

$$\|zv\|_2 = \||z||v|\|_2 \leq \|z\|_2^{1-1/(n-1)} \|v^{n-1}z\|_2^{1/(n-1)}.$$

Combining these inequalities,

$$\|zv\|_2 \leq \|z\|_2^{1-1/(n-1)} \frac{1}{(2nM_{\|v\|})^{1/(p-2)}} H_v(\xi_0)^{1/(p-2)},$$

which concludes the proof. □

Let us return to the convex function

$$f(s) = d_p(1, \delta(s))^p = \tau([\log(e^v e^{sz})]^p).$$

By the Baker–Campbell–Hausdorff formula,

$$\begin{aligned} f(s) &= \tau\left(\left[\sum_{n\geq 0} C_n(v, sz)\right]^p\right) \\ &= \tau\left(\left[v + s\left(z + \frac{1}{2}vz - \frac{1}{2}zv + \frac{1}{12}v^2z - \frac{1}{6}vzv + \frac{1}{12}zv^2\right)\right.\right. \\ &\quad + s^2\left(\frac{1}{12}vz^2 + \frac{1}{12}z^2v - \frac{1}{6}zvz + \frac{1}{24}v^2z^2 - \frac{1}{12}vzvz - \frac{1}{24}z^2v^2 + \frac{1}{12}zvzv\right) \\ &\quad \left.\left. + \cdots\right]^p\right). \end{aligned} \tag{5.3}$$

Remark 5.1. The Taylor series of $f(s)$ can be arranged as follows,

$$\begin{aligned} f(s) &= (-1)^n \tau(v^p) + s^p(-1)^n \tau(z^p) + \sum_{k=1}^{\infty} Q_k(v,z) s^k \\ &= |\tau(v^p)| + s^p|\tau(z^p)| + \sum_{k=1}^{\infty} Q_k(v,z) s^k, \end{aligned}$$

where $Q_k(v,z)$ is a series of monomials in v and z, such that all terms contain either the product vz or vz. This is apparent from 5.3.

Theorem 5.1. *Again suppose $n \geq 2$. There exist constants $C = C(\|v\|, \|z\|) > 0$, $R > 0$, such that if $\|v\| < R$, $\|z\| < R$, then for $k \geq 1$*

$$|Q_k(v,z)| \leq C\|z\|_2^{1-1/(n-1)} f''(0)^{1/(p-2)}$$

In particular, if $f''(0) = 0$, then $Q_k(v,z) = 0$ for $k \geq 1$.

Proof. The series $\log(e^x e^y) = \sum_{k \geq 1} C_k(x,y)$ has an alternate expression given by Dinkin (see for instance [10])

$$\log(e^x e^y) = \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \sum \frac{x^{r_1} y^{s_1} \cdots x^{r_k} y^{s_k}}{r_1! s_1! \cdots r_k! s_k!},$$

where the second sum is taken over all pairs of k-tuples $(r_1, \ldots, r_k)$ and $(s_1, \ldots, s_k)$ such that $r_i + s_i > 0$. This series can be majorized by (see [3])

$$\sum_{k \geq 1} \frac{1}{k} \sum \frac{\|x\|^{r_1} \|y\|^{s_1} \cdots \|x\|^{r_k} \|y\|^{s_k}}{r_1! s_1! \cdots r_k! s_k!} \leq \sum_{k \geq 1} \frac{(e^{\|x\|+\|y\|} - 1)^k}{k},$$

which converges if $\|x\| + \|y\| < \log(2)$. By the remark above, each monomial in the expression of $Q_k(v,z)$ has at least a factor vz or zv. If, for instance, one of these terms is of the form $A(v,z)vzB(v,z)$, then

$$\|A(v,z)vzB(v,z)\|_2 \leq \|A(v,z)\| \|B(v,z)\| \|vz\|_2,$$

with an analogous expression if a factor zv appears (note that $\|vz\|_2 = \|zv\|_2$ because $z, v \in \mathcal{A}_{ah}$). The expressions $Q_k(v,z)$ are the traces of the series of such monomials, thus $|Q_k(v,z)|$ can be bounded by the sum of the 2-norms of these monomials, and therefore

$$|Q_k(v,z)| \leq \|vz\|_2 S(\|v\|, \|z\|),$$

where S is a series of non-negative terms, which are obtained from the terms of the series

$$\sum_{k \geq 1} \frac{(e^{\|x\|+\|y\|} - 1)^k}{k}$$

after factoring out $\|v\|\|z\|$. Therefore it is convergent if $\|v\| + \|z\| < \log(2)$. Thus there exists constants $R > 0$, $M > 0$ such that if $\|v\| < R$ and $\|z\| < R$, then $|Q_k(v, z)| \leq M\|vz\|_2$. The theorem follows using Lemmas 5.1 and 5.4. $\square$

Remark 5.2. Note that if one restricts to $k \geq 3$, the series $S(\|v\|, \|z\|)$ vanishes at the origin. Indeed, after taking the factor zv or vz, at least z remains in every term of the series of $\log(e^v e^{sz})$.

Remark 5.3. The bound R on the norms of v and z can be chosen to be $\frac{1}{6}$. Indeed, this ensures that:

(1) The unitaries $u_0 = 1$, $u_1 = e^v$ and $u_2 = e^z$ verify $\|u_i - u_j\| \leq \frac{1}{2}\sqrt{2 - \sqrt{2}}$. Indeed, $\|e^v - 1\| = 2\sin(\frac{\|v\|}{2})$ which is less than half $\frac{1}{2}\sqrt{2-\sqrt{2}}$ for $\|v\| < \frac{1}{6}$, and analogously for z. Thus $\|e^z - e^v\| \leq \|e^z - 1\| + \|1 - e^v\| < \frac{1}{2}\sqrt{2 - \sqrt{2}}$. Therefore the the map f is convex by virtue of Theorem 4.1.
(2) In Lemma 5.4 $M_{\|v\|} = 1 - 2\|v\| - \frac{2}{3}\|v\|^2 - \frac{2}{3}\|v\|^3 > 0$.
(3) Also $\|v\| + \|z\| < \frac{1}{3} < \log(2)$.

Let us prove now the following convexity condition, which we shall call *p-convexity*.

Theorem 5.2. *Let $v, z \in \mathcal{A}_{ah}$ with $\|v\| < 1/6$ and $\|z\| < 1/6$. Then there exists a positive constant $\epsilon = \epsilon(\|zv\|_2, \|z\|_2)$ such that the function $f(s) = d_p(1, \delta(s))^p$ ($\delta(s) = e^v e^{sz}$) verifies*

$$\frac{1}{s^p}\{f(s) - f(0) - f'(0)s\} \geq |\tau(z^p)|$$

for all s with $|s| < \epsilon$. If $zv = 0$, this inequality holds for all $s \in \mathbb{R}$.

Proof. Recall the Taylor expansion of f.

$$f(s) - f(0) - f'(0)s = \frac{f''(0)}{2}s^2 + (-1)^n s^p \tau(z^p) + \sum_{k\geq 3} Q_k(v, z)s^k.$$

If $f''(0) = 0$, by 5.1, $Q_k(v, z) = 0$ and thus $f(s) = \tau(v^p) + s^p\tau(z^p)$, and the result follows. If $f''(0) > 0$, and again by 5.1, there exists constant $C > 0$ (note that we have fixed $\|v\|, \|z\|$) such that

$$\sum_{k=3}^{p} Q_k(v, z)s^k \geq -C\|z\|^{1-1/(n-1)} f''(0)^{1/(p-2)} \sum_{k=3}^{p} |s|^k.$$

Thus, taking $|s| < 1/2$,

$$\frac{1}{s^p}\{f(s) - f(0) - f'(0)s\} \geq (-1)^n\tau(z^p) + \frac{f''(0)}{2s^{p-2}} - \frac{C}{2|s|^{p-3}} f''(0)^{1/(p-2)}$$

$$= (-1)^{p/2}\tau(z^p) + \frac{f''(0)^{1/(p-2)}}{s^{p-2}}\left\{f''(0)^{1-1/(p-2)} - \frac{C}{2}\|z\|_2^{1-1/(n-1)}|s|\right\}.$$

Therefore to prove our result, it suffices to show that

$$f''(0)^{1-1/(p-2)} - \frac{C}{2}\|z\|_2^{1-1/(n-1)}|s| \geq 0$$

if $|s| < \epsilon$, for a given $\epsilon > 0$. This ϵ clearly exists, it only remains to see that it can be chosen to depend on $\|zv\|_2$ and $\|z\|_2$. Suppose first that $zv \neq 0$. Recall from (5.1) that if $\|v\| \leq \alpha < \pi/2$, then $f''(0) \geq H_v(\xi_0)\cos(\alpha)$. Here we choose $\alpha = 1/6$. Combining inequalities 5.1 and 5.2 in the proof of Lemma (5.4), one obtains

$$H_v(\xi_0) \geq p\|zv^{n-1}\|_2^2 M_{\|v\|}.$$

Therefore, in our case ($\|v\| \leq 1/6$),

$$f''(0) \geq \cos(1/6)M_{1/6}\ p\|zv^{n-1}\|_2^2.$$

Next recall that (Lemma (5.3))

$$\|zv^{n-1}\|_2^{1/(n-1)} \geq \frac{\|zv\|_2}{\|z\|_2^{1-1/(n-1)}}.$$

Thus there is a constant $D > 0$ such that

$$f''(0) \geq D\frac{\|zv\|_2^{n-1}}{\|z\|_2^{n-2}}.$$

Take

$$\epsilon = \frac{2}{C}\left(D\frac{\|zv\|_2^{n-1}}{\|z\|_2^{n-2}}\right)^{1-1/(p-2)} \frac{1}{\|z\|_2^{1-1/(n-1)}}.$$

If $|s| < \epsilon$, then a straightforward use of the above inequalities proves that

$$f''(0)^{1-1/(p-2)} - \frac{C}{2}\|z\|_2^{1-1/(n-1)}|s| \geq 0.$$

If $zv = 0$, $f(s) = (-1)^n\tau(v^p) + s^p(-1)^n\tau(z^p)$, and the inequality holds for all s. $\square$

Acknowledgments

The first named author is supported by the PIP 5690-CONICET grant.

References

[1] E. Andruchow, Short geodesics of unitaries in the L^2-metric, *Can. Math. Bull.* **48**(3) (2005) 340–354.

[2] E. Andruchow and L. Recht, Grassmannians of a finite algebra in the strong operator topology, *Int. J. Math.* **17** (2006) 477–491.

[3] P. Cartier, Structure topologique des groupes de Liegénérayx, Séminaire "Sophus Lie" 22, Ecole Normale Supérieure, 1954/55.

[4] C. E. Durán, L. E. Mata-Lorenzo and L. Recht, Natural variational problems in the Grassmann manifold of a C*-algebra with trace, *Adv. Math.* **154** (2000) 196–228.

[5] C. E. Durán, L. E. Mata-Lorenzo and L. Recht, Metric geometry in homogeneous spaces of the unitary group of a C*-algebra. I. Minimal curves, *Adv. Math.* **184** (2004) 342–366.

[6] C. E. Durán, L. E. Mata-Lorenzo and L. Recht, Metric geometry in homogeneous spaces of the unitary group of a C*-algebra. II. Geodesics joining fixed endpoints, *Integral Equations Operator Theory* **53** (2005) 33–50.

[7] G. H. Hardy, J. E. Littlewood and G. Pòlya, *Inequalities*. Reprint of the 1952 edition. Cambridge Mathematical Library (Cambridge University Press, Cambridge, 1978).

[8] J. Jost, Nonpositive Curvature, Geometric and Analytic Aspects, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 1997.

[9] L. E. Mata-Lorenzo and L. Recht, Convexity properties of $\mathrm{Tr}[(a^*a)^n]$, *Linear Alg. Appl.* **315** (2000) 25–38.

[10] W. Rossmann, *Lie Groups: An Introduction Through Linear Groups* (Oxford University Press, Oxford, 2002).

[11] H. Upmeier, *Symmetric Banach Manifolds and Jordan C^*-Algebras*, North-Holland Mathematics Studies, Vol. 104; Notas de Matemtica [Mathematical Notes], Vol. 96 (North-Holland Publishing Co., Amsterdam, 1985).

[12] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*, Prentice-Hall Series in Modern Analysis (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974).