ILLINOIS JOURNAL OF MATHEMATICS
Volume 38, Number 1, Spring 1994

CONVEXITY OF THE GEODESIC DISTANCE ON SPACES OF POSITIVE OPERATORS

G. CORACH, H. PORTA AND L. RECHT

Let A be a C^*-algebra with 1 and denote by A^+ the set of positive invertible elements of A. The set A^+ being open in $A^s = \{a \in A;\ a^* = a\}$ it has a C^∞ structure and we can identify TA_a^+ with A^s for each $a \in A^+$. We use G to denote the group of invertible elements of A. Notice that G operates on the left on A^+ by the rule

$$L_g a = (g^*)^{-1} a g^{-1} \quad (g \in G, a \in A^+).$$

This action allows us to introduce a natural reductive homogeneous space structure in the sense of [8] (for details see [2], [3], [4]).

The corresponding connection—which is preserved by the group action—has covariant derivative

$$\frac{DX}{dt} = \frac{dX}{dt} - \frac{1}{2}(\dot{\gamma}\gamma^{-1}X + X\gamma^{-1}\dot{\gamma})$$

where X is a tangent field on A^+ along the curve γ and exponential

$$\exp_a X = e^{Xa^{-1}/2} a e^{a^{-1}X/2}, \quad a \in A^+, X \in TA_a^+.$$

The curvature tensor has the formula

$$R(X,Y)Z = -\tfrac{1}{4}a[[a^{-1}X, a^{-1}Y], a^{-1}Z]$$

for $X, Y, Z \in TA_a^+$. The manifold A^+ has also a natural Finsler structure given by

$$\|X\|_a = \|a^{-1/2}Xa^{-1/2}\| \text{ for } X \in TA_a^+$$

and the group G operates by isometries for this Finsler metric.

THEOREM 1. *If $J(t)$ is a Jacobi field along the geodesic $\gamma(t)$ in A^+ then $\|J(t)\|_{\gamma(t)}$ is a convex function of $t \in \mathbf{R}$.*

Received February 11, 1992
1991 Mathematics Subject Classification. Primary 53C22; Secondary 58B20, 52AD5.

Proof. The method of proof is based on a similar strategy used in [4]. By definition $J(t)$ satisfies the equation

$$\frac{D^2J}{dt^2} + R(J,V)V = 0 \tag{1}$$

where $V(t) = \dot{\gamma}(t)$.

Notice that by the invariance of the connection and the metric under the action of G we may assume that $\gamma(t) = e^{tX}$ is a geodesic starting at $\gamma(0) = 1 \in A$, where $X \in A^s$. Then for the field $K(t) = e^{-tX/2}J(t)e^{-tX/2}$ the differential equation (1) changes into

$$4\ddot{K} = KX^2 + X^2K - 2XKX, \tag{2}$$

(where the dots indicate ordinary derivative with respect to t). Since the group G acts by isometries, we have $\|J(t)\|_{\gamma(t)} = \|\gamma(t)^{-1/2}J(t)\gamma(t)^{-1/2}\| = \|K(t)\|$. Thus the proof reduces to showing that for any solution $K(t)$ of (2) the function $t \to \|K(t)\|$ is convex in $t \in \mathbf{R}$, where the norm is the *ordinary norm* in the C^* algebra A. So fix $u < v \in \mathbf{R}$ and let t satisfy $u \leq t \leq v$. We will prove that

$$\|K(t)\| \leq \frac{v-t}{v-u}\|K(u)\| + \frac{t-u}{v-u}\|K(v)\|. \tag{3}$$

Consider first the case where the selfadjoint element $X \in A$ has the form

$$X = \sum_{i=1}^{n} \lambda_i p_i \tag{4}$$

with $\lambda_1, \lambda_2, \dots \lambda_n$ real numbers and $p_1, p_2, \dots p_n$ selfadjoint elements of A satisfying $p_i p_j = 0$ for $i \neq j$ and $p_1 + p_2 + \cdots + p_n = 1$.

Suppose that A is faithfully represented in a Hilbert space $\mathcal{H}$. For fixed $x \in A$ decompose $x \in \mathcal{H}$ as $x = \sum_{i=1}^{n} \xi_i x_i$ where x_i is a unit vector in the range of p_i and the ξ_i are appropriate scalars. Define next the matrix $k(t) = (k_{ij}(t))$ by $k_{ij}(t) = \langle K(t)x_i, x_j\rangle$ for all t. The differential equation (2) is equivalent to the equations

$$\ddot{k}_{ij}(t) = \delta_{ij}^2 k_{ij}(t) \tag{2ij}$$

where $\delta_{ij} = (\lambda_i - \lambda_j)/2$.

A simple verification (or Bernoulli's formula) shows that all solutions of $\ddot{f}(t) = c^2 f(t)$ satisfy

$$f(t) = \phi(u,v,c;t)f(u) + \psi(u,v,c;t)f(v)$$

where

$$\phi(u,v,c;t) = \begin{cases} \dfrac{\operatorname{Sinh} c(v-t)}{\operatorname{Sinh} c(v-u)} & \text{for } c \neq 0, \\ \dfrac{(v-t)}{(v-u)} & \text{for } c = 0, \end{cases}$$

$$\psi(u,v,c;t) = \begin{cases} \dfrac{\operatorname{Sinh} c(t-u)}{\operatorname{Sinh} c(v-u)} & \text{for } c \neq 0, \\ \dfrac{(t-u)}{(v-u)} & \text{for } c = 0. \end{cases}$$

Then each $k_{ij}(t)$ satisfies

$$k_{ij}(t) = \phi_{ij}(t)k_{ij}(u) + \psi_{ij}(t)k_{ij}(v)$$

where $\phi_{ij}(t) = \phi(u, v, \delta_{ij}; t)$ and $\psi_{ij}(t) = \psi(u, v, \delta_{ij}; t)$. This can be written in matrix form as

$$k(t) = \Phi(t) \circ k(u) + \Psi(t) \circ k(v)$$

where $\Phi(t) = \{\phi_{ij}(t)\}$ and $\Psi(t) = \{\psi_{ij}(t)\}$,and the symbol $\circ$ denotes the Schur product $\{a_{ij}\} \circ \{b_{ij}\} = \{a_{ij}b_{ij}\}$ of matrices. It follows that

$$\|k(t)\| \leq \|\Phi(t) \circ k(u)\| + \|\Psi(t) \circ k(v)\|. \tag{5}$$

The final step is to prove the inequalities

$$\|\Phi(t) \circ k(u)\| \leq \frac{v-t}{v-u}\|k(u)\|,$$
$$\|\Psi(t) \circ k(v)\| \leq \frac{t-u}{v-u}\|k(v)\|. \tag{6}$$

Notice that both $\Phi(t)$ and $\Psi(t)$ are positive semidefinite. This follows from Bochner's theorem [1] applied to $\phi(u, v, c; t)$ and $\psi(u, v, c; t)$ considered as functions of c. In both cases the matrix is of the form $\{F(\lambda_i - \lambda_j)\}$ where $F(c)$ is the Fourier transform of a positive function (see [7], formula 1.9.14, page 31).

Next we apply a theorem of Davis (see [6] and the generalization in [9]) according to which for $n \times n$-matrices A and P with P positive semidefinite we have

$$\|P \circ A\| \leq \Big(\max_{1 \leq i \leq n} P_{ii}\Big)\|A\|.$$

Taking $P = \Phi(t)$ and $P = \Psi(t)$ we get inequalities (6). Using now (5) and (6) we also get

$$\|k(t)\| \leq \frac{v-t}{v-u}\|k(u)\| + \frac{t-u}{v-u}\|k(v)\|. \tag{7}$$

Since the element x and the representation space $\mathscr{H}$ were not specified, we may assume without loss of generality that for a given t between u and v we have $\|K(t)x\| = |\langle K(t)x, x\rangle|$. Then writing $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ we conclude that

$$\begin{aligned} |\langle k(t)\xi, \xi\rangle| &= |\langle K(t)x, x\rangle| = \|K(t)\| \\ |\langle k(u)\xi, \xi\rangle| &= |\langle K(u)x, x\rangle| \leq \|K(t)\| \\ |\langle k(v)\xi, \xi\rangle| &= |\langle K(v)x, x\rangle| \leq \|K(t)\| \end{aligned}$$

and then (3) follows from (7) for X of the special form (4).

Let us go then to the general case—when X is an arbitrary selfadjoint element of A. The spectral theorem allows us to approximate X (in operator norm) by elements of the form (4). From the well-possedness of problem (2) we conclude that $(t, X) \to K(t)$ is norm continuous, and the inequality (3) for arbitrary X follows from the same inequality for X of the form (4). This completes the proof of Theorem 1.

For $a, b \in A^+$ let $\operatorname{dist}(a, b)$ denote the geodesic distance from a to b in the Finsler metric $\|X\|_a$ of A. It is not hard to prove (using the invariance of the metric) that

$$\operatorname{dist}(a, b) = \|\ln(a^{-1/2}ba^{-1/2})\|. \tag{8}$$

THEOREM 2. *If $\gamma(t)$ and $\delta(t)$ are geodesics in A^+ then $t \to \operatorname{dist}(\gamma(t), \delta(t))$ is a convex function of $t \in \mathbf{R}$.*

Proof. Suppose the geodesics $\gamma(t)$ and $\delta(t)$ are defined for $u \leq t \leq v$. Define $h(s, t)$ by the properties:

(a) the function $s \to h(s, u)$, $0 \leq s \leq 1$ is the geodesic joining $\gamma(u)$ and $\delta(u)$;

(b) the function $s \to h(s, v)$, $0 \leq s \leq 1$ is the geodesic joining $\gamma(v)$ and $\delta(v)$;

(c) for each s, the function $t \to h(s, t)$, $u \leq s \leq v$ is the geodesic joining $h(s, u)$ and $h(s, v)$.

In particular $h(0, t) = \gamma(t)$ and $h(1, t) = \delta(t)$. Define also $J(s, t) = \partial h(s, t)/\partial s$. Then, for each $s, t \to J(s, t)$ is a Jacobi field along the geodesic

$t \to h(s,t)$. Finally define

$$f(t) = \int_0^1 \| J(s,t) \|_{h(s,t)}\, ds.$$

From Theorem 1, $t \to \|J(s,t)\|$ is convex for each s. Hence $t \to f(t)$ is also convex for $u \le t \le v$. But $f(u) = \int_0^1 \|J(s,u)\|_{h(s,u)}\, ds$ is the length of the geodesic $s \to h(s,u)$ and therefore $f(u) = \operatorname{dist}(\gamma(u), \delta(u))$. Similarly, $f(v) = \operatorname{dist}(\gamma(v), \delta(v))$. Now for $u \le t \le v$, the value $f(t) = \int_0^1 \|J(s,t)\|_{h(s,t)}\, ds$ is the length of the curve $s \to h(s,t)$ joining $\gamma(t)$ and $\delta(t)$ and then we have $\operatorname{dist}(\gamma(v), \delta(v)) \le f(t)$. Convexity of $\operatorname{dist}(\gamma(v), \delta(v))$ follows and Theorem 2 is proved.

COROLLARY 2.1. *For any fixed $y \in A^+$ the function $f\colon A^+ \to \mathbf{R}$, $f(x) = \operatorname{dist}(x, y)$ is |convex in the geometric sense", that is, each geodesic $\gamma(t)$ satisfies*

$$f(\gamma(t)) \le (1-t) f(\gamma(0)) + t f(\gamma(1)).$$

In particular geodesic spheres are convex sets.

Proof. Take $\delta(t) = y$ for all t and apply Theorem 2.

COROLLARY 2.2. *For any a_0, a_1, b_0, and b_1 in A^+ we have*

$$\begin{aligned} &\left\| \left(a_0^{1/2}\left(a_0^{-1/2} a_1 a_0^{-1/2}\right)^t a_0^{1/2}\right)^{1/2} \left(b_0^{1/2}\left(b_0^{-1/2} b_1 b_0^{-1/2}\right)^t b_0^{1/2}\right)^{1/2} \right\| \\ &\qquad \le \|a_0^{1/2} b_0^{1/2}\|^{1-t} \|a_1^{1/2} b_1^{1/2}\|^t. \end{aligned} \tag{9}$$

Proof. Take two geodesics $\gamma(t)$ and $\delta(t)$ and write them as

$$\begin{aligned} \gamma(t) &= a_0^{1/2}\left(a_0^{-1/2} a_1 a_0^{-1/2}\right)^t a_0^{1/2}, \\ \delta(t) &= b_0^{1/2}\left(b_0^{-1/2} b_1 b_0^{-1/2}\right)^t b_0^{1/2} \end{aligned}$$

where $a_0 = \gamma(0)$, $a_1 = \gamma(1)$, $b_0 = \delta(0)$, $b_1 = \delta(1)$. Then for each $0 \le t \le 1$ we have, by convexity,

$$\operatorname{dist}(\gamma(t), \delta(t)) \le (1-t) \operatorname{dist}(a_0, b_0) + t \operatorname{dist}(a_1, b_1)$$

or

$$\begin{aligned} &\left\| \ln\left(\gamma(t)^{-1/2} \delta(t) \gamma(t)^{-1/2}\right) \right\| \\ &\qquad \le (1-t) \left\| \ln\left(a_0^{-1/2} b_0 a_0^{-1/2}\right) \right\| + t \left\| \ln\left(a_1^{-1/2} b_1 a_1^{-1/2}\right) \right\|. \end{aligned}$$

Next we apply this formula to the geodesics $\gamma(t)$ and $k\delta(t)$ where $k > 0$. By choosing k large enough we can assume that

$$\gamma(t)^{-1/2}(k\delta(t))\gamma(t)^{-1/2} > 1$$
$$a_0^{-1/2}(kb_0)a_0^{-1/2} > 1$$
$$a_1^{-1/2}(kb_1)a_1^{-1/2} > 1$$

and therefore using $\|\ln x\| = \ln\|x\|$ for $x > 1$ and canceling out k, the last inequality for norms becomes

$$\left\|\gamma(t)^{-1/2}\delta(t)\gamma(t)^{-1/2}\right\| \leq \|a_0^{-1/2}b_0a_0^{-1/2}\|^{1-t}\|a_1^{-1/2}b_1a_1^{-1/2}\|^t.$$

Notice that $\gamma(t)^{-1}$ is also a geodesic so that the last formula gives also:

$$\left\|\gamma(t)^{1/2}\delta(t)\gamma(t)^{1/2}\right\| \leq \|a_0^{1/2}b_0a_0^{1/2}\|^{1-t}\|a_1^{1/2}b_1a_1^{1/2}\|^t$$

or equivalently

$$\left\|\gamma(t)^{1/2}\delta(t)^{1/2}\right\| \leq \|a_0^{1/2}b_0^{1/2}\|^{1-t}\|a_1^{1/2}b_1^{1/2}\|^t.$$

which is another way to write (9).

This inequality has many variations. For example, replacing a_i by a_i^2 and b_i by b_i^2 and using the definition of the geodesics, we get

$$\left\|\left(a_0(a_0^{-1}a_1^2a_0^{-1})^ta_0\right)^{-1/2}\left(b_0(b_0^{-1}b_1^2b_0^{-1})^tb_0\right)^{-1/2}\right\| \leq \|a_0b_0\|^{1-t}\|a_1b_1\|^t$$

or using $|z| = (zz^*)^{1/2}$:

$$\left\|\,\left|a_0(a_0^{-1}a_1^2a_0^{-1})^{t/2}\right|\left|b_0(b_0^{-1}b_1^2b_0^{-1})^{1/2}\right|\right\| \leq \|a_0b_0\|^{1-t}\|a_1b_1\|^t.$$

As special cases of (9) we can also get $\|ab^ta\| \leq \|aba\|^t$ and $\|a^tb^t\| \leq \|ab\|^t$ for any $a, b \in A^+$ and $0 \leq t \leq 1$.

THEOREM 3 (see [3]). *The exponential function in A^+ increases distances.*

Proof. By invariance it suffices to show that the exponential function increases distances at the identity $1 \in A^+$. Consider two geodesics of the form $\gamma(t) = e^{tX}$ and $\delta(t) = e^{tY}$. Then according to Theorem 2 the function

$$f(t) = \operatorname{dist}(\gamma(t), \delta(t)) = \|\ln(e^{-tX/2}e^{tY}e^{-tX/2})\|$$

is convex. Since $f(0) = 0$ this implies that $f(t)/t \leq f(1)$ for each $0 < t \leq 1$. Taking limits we have $\lim_{t \to 0} f(t)/t \leq f(1)$.

Observe next that $\ln x$ can be approximated on any interval $[x_0, x_1]$ with $0 < x_0 < x_1$ uniformly in the C^1 sense by polynomials $p_n(x)$. In particular $\lim_{n \to \infty} p_n(x) = \ln x$ and $\lim_{n \to \infty} p_n'(x) = 1/x$. Then

$$\begin{aligned}
&\lim_{t \to 0} \frac{1}{t} \ln(e^{-tX/2} e^{tY} e^{-tX/2}) \\
&\quad = \lim_{n \to \infty} \lim_{t \to 0} \frac{1}{t} p_n(e^{-tX/2} e^{tY} e^{-tX/2}) \\
&\quad = \lim_{n \to \infty} \frac{d}{dt} p_n(e^{-tX/2} e^{tY} e^{-tX/2}) \Big|_{t=0} = Y - X
\end{aligned}$$

(the last inequality is justified below). Now from this equality and convexity we conclude that $f(t) \geq t\|Y - X\|$ and this means that

$$\operatorname{dist}(\exp_a(tX), \exp_a(tY)) \geq t\|Y - X\| \quad \text{for all } a \in A^+ \quad \text{and all } X, Y \in TA_a^+.$$

To finish the proof write the polynomials p_n explicitly as $p_n(x) = \Sigma r_{n,k} x^k$. Then

$$\begin{aligned}
&\frac{d}{dt} \ln(e^{-tX/2} e^{tY} e^{-tX/2})|_{t=0} \\
&\quad = \lim_{n \to \infty} \frac{d}{dt} p_n(e^{-tX/2} e^{tY} e^{-tX/2}) \Big|_{t=0} \\
&\quad = \lim_{n \to \infty} \sum r_{n,k} \frac{d}{dt} (e^{-tX/2} e^{tY} e^{-tX/2})^k \Big|_{t=0} \\
&\quad = \lim_{n \to \infty} \sum r_{n,k} (Y - X)^k = \lim_{n \to \infty} p_n'(1)(Y - X) = (Y - X).
\end{aligned}$$

As observed in [3] this property of the exponential is equivalent to Segal's inequality ($\|e^{X+Y}\| \leq \|e^X e^{tY}\|$ for X, Y selfadjoint) which is therefore another consequence of the convexity of the distance function in A^+.

References

1. S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
2. G. Corach, H. Porta and L. Recht, *The geometry of the space of self-adjoint invertible elements of a C^*-algebra*, to appear (preprint form in Trabajos de Matemática #149, Instituto Argentino de Matemática.)
3. ______, *A geometric interpretation of Segal's inequality*, Proc. Amer. Math. Soc., **115** (1992), 229–231.
4. ______, *Jacobi fields on spaces of positive operators*, Linear Algebra Appl., to appear.
5. ______, *Geodesics and operator means in the space of positive operators*, Internat. J. Math., to appear.

6. C. DAVIS, *On the norm of the Schur product operation*, Numer. Math. **4** (1962), 343–344.
7. A. ERDELYI ET AL., *Tables of integral transforms* (The Bateman manuscript), Volume 1, McGraw-Hill, New York, 1954.
8. M. KOBAYASHI and K. NOMIZU, *Foundations of differential geometry*, Interscience, New York, 1969.
9. M. WALTER, *On the norm of the Schur product*, Linear Algebra Appl. **79** (1986), 209–213.

INSTITUTO ARGENTINO DE MATEMÁTICA
BUENOS AIRES, ARGENTINA

UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

UNIVERSIDAD SIMÓN BOLÍVAR
CARACAS, VENEZUELA