Nonpositive curvature: A geometrical approach to Hilbert–Schmidt operators

Gabriel Larotonda *

Instituto de Ciencias, Universidad Nacional de General Sarmiento,
JM Gutierrez 1150 (1613) Los Polvorines, Buenos Aires, Argentina

Received 10 June 2007

Communicated by W.T. Meyer

Abstract

We give a Riemannian structure to the set Σ of positive invertible unitized Hilbert–Schmidt operators, by means of the trace inner product. This metric makes of Σ a nonpositively curved, simply connected and metrically complete Hilbert manifold. The manifold Σ is a universal model for symmetric spaces of the noncompact type: any such space can be isometrically embedded into Σ. We give an intrinsic algebraic characterization of convex closed submanifolds M. We study the group of isometries of such submanifolds: we prove that G_M, the Banach–Lie group generated by M, acts isometrically and transitively on M. Moreover, G_M admits a polar decomposition relative to M, namely $G_M \simeq M \times K$ as Hilbert manifolds (here K is the isotropy of $p = 1$ for the action $I_g : p \mapsto gpg^*$), and also $G_M/K \simeq M$ so M is an homogeneous space. We obtain several decomposition theorems by means of geodesically convex submanifolds M. These decompositions are obtained *via* a nonlinear but analytic orthogonal projection $\Pi_M : \Sigma \to M$, a map which is a contraction for the geodesic distance. As a byproduct, we prove the isomorphism $NM \simeq \Sigma$ (here NM stands for the normal bundle of a convex closed submanifold M). Writing down the factorizations for fixed e^a, we obtain $e^a = e^x e^v e^x$ with $e^x \in M$ and v orthogonal to M at $p = 1$. As a corollary we obtain decompositions for the full group of invertible elements $G \simeq M \times \exp(T_1 M^\perp) \times K$.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

The aim of this paper is to relate the algebraic and spectral properties of the Banach algebra of unitized Hilbert–Schmidt operators, with the metric and geometrical properties of an underlying manifold Σ. This is a paper on *applied* nonpositively curved geometry because we first show how the familiar properties of the operator algebra translate into geometrical notions, and then we use the tools of geometry in order to prove new results concerning the operator algebra.

* Tel./fax: +54 011 44697501.
  E-mail address: glaroton@ungs.edu.ar.

In this paper we study the cone of positive invertible Hilbert–Schmidt operators (extended by the scalar operators) on a separable infinite dimensional Hilbert space H. The metric in the tangent space at the identity is given by the trace of the algebra. The local structure induced by the metric is smooth and quadratic; it can be situated in the context of the theory of infinite dimensional Riemann–Hilbert manifolds of nonpositive curvature (cf. Cartan–Hadamard manifolds, as introduced by Lang [18], McAlpin [21], Grossman [14] and others). It is then a paper on Riemannian geometry. On the other hand, since the manifold Σ is clearly not locally compact, some of the standard results for Hadamard manifolds require a different approach. The geometry is then related to the geometry of the metric spaces in the sense of Aleksandrov [5]. It turns out that the notion of convexity (together with the fact that Σ is a simply connected and globally nonpositively curved geodesic length space) plays a key role in our constructions. It is then a paper on *metric* geometry.

Through the years, several authors have studied the relationship of geometry and algebra in sets of positive operators, with different approaches that led to a variety of results. In his 1955's paper [22], G.D. Mostow gave a Riemannian structure to the set M_n^+ of positive invertible matrices; the induced metric makes of M_n^+ a nonpositively curved symmetric space. Mostow showed that the algebraic concept behind the notion of convexity is that of a *Lie triple system*, which is basically the real part of a given involutive Lie algebra $\mathfrak{g}$. The geometry of bounded positive operators in an infinite dimensional Hilbert space was studied by G. Corach, H. Porta and L. Recht [10,12,24] among others, using functional analysis techniques. This area of research is currently very active (see [8,9] for a list of references).

1.1. Main results

In this paper we study the geometry of a Hilbert manifold Σ which is modeled on the operator algebra $\mathcal{H}_{\mathbb{C}}$ of unitized Hilbert–Schmidt operators. In Section 2 we introduce the objects involved and prove some elementary results. The manifold Σ is the set of positive invertible operators of $\mathcal{H}_{\mathbb{C}}$. Let $\mathcal{H}_{\mathbb{C}}^{\bullet}$ be the classical Banach–Lie group of invertible (unitized) Hilbert–Schmidt operators [16]. The manifold Σ has a natural $\mathcal{H}_{\mathbb{C}}^{\bullet}$-invariant metric $\langle x, y \rangle_p = \langle xp^{-1}, p^{-1}y \rangle_2$, which makes it nonpositively curved (we define $\langle \alpha + a, \beta + b \rangle_2 = \alpha\overline{\beta} + \mathrm{tr}(b^*a)$ whenever $\alpha, \beta \in \mathbb{C}$ and a, b are Hilbert–Schmidt operators). Let e^x and $\exp(x)$ stand for the usual analytic exponential, i.e., $\exp(x) = \sum_{n \geqslant 0} \frac{x^n}{n!}$. This map is injective when restricted to $\mathcal{H}_{\mathbb{R}}$, the set of self-adjoint operators. Let $\ln(p)$ stand for its real analytic inverse. We have $\exp(\mathcal{H}_{\mathbb{R}}) = \Sigma \subset \mathcal{H}_{\mathbb{R}}$, and the exponential map induces a diffeomorphism onto its image, so we identify the tangent space at any point of the manifold Σ with the set of self-adjoint operators $\mathcal{H}_{\mathbb{R}}$, namely $T_p\Sigma \simeq \mathcal{H}_{\mathbb{R}}$ for any $p \in \Sigma$. In Section 3 we prove

Theorem A. *For $p, q \in \Sigma$, the geodesic obtained from Euler's equation by solving Dirichlet's problem is the smooth curve $\gamma_{pq}(t) = p^{1/2}(p^{-1/2}qp^{-1/2})^t p^{1/2}$, hence*

$$\mathrm{Exp}_p(v) = p^{\frac{1}{2}} \exp(p^{-\frac{1}{2}} v \, p^{-\frac{1}{2}}) p^{\frac{1}{2}}$$

is the Riemannian exponential of Σ, for any $v \in T_p\Sigma$. Both $\mathrm{Exp}_p : T_p\Sigma \to \Sigma$ and its differential map $d(\mathrm{Exp}_p)_v : T_p\Sigma \to T_{\mathrm{Exp}_p(v)}\Sigma$ are C^ω-isomorphisms for any $p \in \Sigma$, $v \in T_p\Sigma$. The curve γ_{pq} is the shortest piecewise smooth path joining p to q, hence

$$\mathrm{dist}(p, q) = \left\| \ln(p^{\frac{1}{2}} q^{-1} p^{\frac{1}{2}}) \right\|_2$$

is the distance in Σ induced by the Riemannian metric. The metric space (Σ, dist) is complete, and it is globally nonpositively curved.

The curve obtained via Calderón's method of complex interpolation [7] between the quadratic norms $\| \cdot \|_p$ and $\| \cdot \|_q$ is exactly the short geodesic in Σ joining p to q (the proof of [1] can be adapted almost verbatim).

In [14], N. Grossman proves that the inequality

$$\left\| d(\mathrm{Exp}_p)_v(w) \right\|_p \geqslant \|w\|_{\mathrm{Exp}_p(v)} \tag{1}$$

leads to the minimality of geodesics in a simply connected, complete Hilbert manifold. This approach is also carried out by McAlpin [21]. The following operator inequality involving the differential of the usual exponential map

$$\left\| \mathrm{e}^{-x/2} \, d\exp_x(y)\mathrm{e}^{-x/2} \right\|_2 \geqslant \|y\|_2 \tag{2}$$

is the translation to our context of the inequality (1) above. The convexity of Jacobi fields can be deduced from the nonpositiveness of the sectional curvature, hence the proof of Eq. (2) stems in our context from the Cauchy–Schwarz inequality for the trace inner product. We follow the exposition of Lang [19] on this subject. On the other hand, (2) can be proved with a direct computation [6]. With this approach the metric completeness of the tangent spaces is not relevant: in Theorem 3.1 of [2], the authors prove the minimizing property of the geodesics in a non complete manifold. The inequality above, in our context, can be also interpreted as the Hyperbolic Cosine Law (see Corollary 3.12)

$$a^2 \geqslant b^2 + c^2 - 2bc \cos(\alpha).$$

Here a, b, c are the lengths of the sides of any geodesic triangle in Σ, and α is the angle opposite to a. From this inequality also follows that the sum of the inner angles of any geodesic triangle in Σ is bounded by π.

If A is a set of operators, we use A^+ to denote the set of positive operators of A; note that $(\mathcal{H}_{\mathbb{C}}^{\bullet})^+ = \Sigma$. In Section 4 we show that a submanifold $M \subset \Sigma$ is geodesically convex if and only if its tangent space at the identity $\mathfrak{m}$ is a Lie triple system. Clearly any such submanifold is nonpositively curved, and Theorem 4.18 states:

Theorem B. *For any geodesically convex, closed submanifold $M = \exp(\mathfrak{m}) \subset \Sigma$ there exists a connected Banach–Lie group $G_M = \langle \exp(\mathfrak{m} \oplus [\mathfrak{m}, \mathfrak{m}]) \rangle \subset \mathcal{H}_{\mathbb{C}}^{\bullet}$ which acts isometrically and transitively on M. Moreover, the polar decomposition of the elements of G_M reduces to M in the sense that $G_M^+ = M$. Let K be the isotropy of 1 for the action; then K is a connected Banach–Lie subgroup of G_M and there is an isomorphism $G_M \simeq M \times K$. In particular any convex submanifold M of Σ is an homogeneous space for a suitable Banach–Lie group, which is an analytical subgroup of $\mathcal{H}_{\mathbb{C}}^{\bullet}$. The submanifold M is flat if and only if $M \equiv G_M$ is an abelian Banach–Lie subgroup of $\mathcal{H}_{\mathbb{C}}^{\bullet}$.*

The existence of smooth polar decompositions for the involutive Banach–Lie groups can be obtained from the general results of Neeb [23, Theorem 5.1]. Neeb introduces the notion of seminegative curvature (SNC) on Banach–Finsler manifolds M, given by the condition of inequality (1) above, plus the condition that $d(\mathrm{Exp}_p)_v$ should be invertible for any $v \in T_p M$ (the metric of M should be invariant under parallel transport along geodesics). Neeb proves (Theorem 1.10 of [23]) that in a connected, geodesically complete manifold with SNC, the exponential map $\mathrm{Exp}_p : T_p M \to M$ is a covering map and M is metrically complete, a result which extends that of Grossman and McAlpin mentioned above to the Banach–Finsler context.

The manifold Σ can be decomposed by means of any convex closed submanifold M. Let NM be the normal bundle of M. In Section 5 we prove

Theorem C. *For any convex closed submanifold $M \subset \Sigma$ there is a nonlinear, real analytic projection $\Pi_M : \Sigma \to M$, which is contractive for the geodesic distance*

$$\mathrm{dist}\big(\Pi_M(p), \Pi_M(q)\big) \leqslant \mathrm{dist}(p, q) \quad \text{for any } p, q \in \Sigma.$$

The point $\Pi_M(p)$ is the (unique) point of M closest to p. It can also be viewed as the unique point in M such that there exists a geodesic through p orthogonal to M at $\Pi_M(p)$. The exponential map $(p, v) \mapsto \mathrm{Exp}_p(v)$ induces an analytic Riemannian isomorphism $NM \simeq \Sigma$.

Since $\Pi_M(p)$ is the point in M closest to p, one can prove the existence of such a point using a metric argument valid in any nonpositively curved geodesic length space [17]. We choose to give a differential-geometry argument here.

In Section 6 we exhibit a decomposition for the submanifold $M = \Delta$ of positive diagonal operators, which is a maximal abelian subalgebra of $\mathcal{H}_{\mathbb{C}}$. This decomposition theorem (Theorem 6.2) takes the form of a factorization $e^a = d e^v d$, where v has null diagonal and d is an invertible diagonal operator. We stress that there is no known algorithm that allows to compute d explicitly (not even if we reduce the problem to 3×3 matrices, that can be thought of as a particular case of the general theory). As a corollary to the decomposition theorems we obtain

Theorem D. *Any invertible operator $g \in \mathcal{H}_{\mathbb{C}}^{\bullet}$ admits a unique polar decomposition relative to a fixed closed convex submanifold $M = \exp(\mathfrak{m})$. Namely $g = e^x e^v u$ where $x \in \mathfrak{m}$, $v \in \mathfrak{m}^{\perp}$ and $u \in \mathcal{U}(\mathcal{H}_{\mathbb{C}})$ is a unitary operator. The map $g \mapsto (e^x, e^v, u)$ is an analytic bijection which gives the isomorphism*

$$\mathcal{H}_{\mathbb{C}}^{\bullet} \simeq M \times \exp(\mathfrak{m}^{\perp}) \times \mathcal{U}(\mathcal{H}_{\mathbb{C}}).$$

This isomorphism generalizes the decomposition of M_n^+ given in [11].

In Section 7 we show that the manifold Σ can be decomposed by means of a foliation $\{\Sigma_\lambda\}_{\lambda>0}$ of totally geodesic submanifolds, namely

$$\Sigma = \dot{\bigcup_{\lambda>0}} \Sigma_\lambda = \dot{\bigcup_{\lambda>0}} \{a + \lambda \in \Sigma,\ a = a^* \text{ a Hilbert–Schmidt operator}\}.$$

There is a Riemannian isomorphism $\Sigma \simeq \Sigma_1 \times \mathbb{R}_{>0}$ induced by the projection Π_{Σ_1} of Theorem C above. As an application, we show a decomposition relative to the algebra M_n^+ of positive invertible $n \times n$ matrices: fix an n-dimensional subspace $S \subset H$, let P_S be the orthogonal projection to S and $Q_S = 1 - P_S$ the orthogonal projection to $S^\perp$. Let $B(S)$ stand for the algebra of bounded linear operators of S. Let $R \in B(S)^+ \simeq M_n^+$, and consider the set

$$\mathfrak{v} = \left\{ \begin{pmatrix} 0 & Y^* \\ Y & X \end{pmatrix}:\ X = X^* \in B(S^\perp) \text{ a Hilbert–Schmidt operator},\ Y \in B(S, S^\perp) \right\}.$$

Let $\mathcal{U}(\mathcal{H}_\mathbb{C})$ be the Banach–Lie subgroup of unitary operators in $\mathcal{H}_\mathbb{C}^\bullet$.

Theorem E. *For any $g \in \mathcal{H}_\mathbb{C}^\bullet$ there is a unique factorization $g = \lambda r e^v u$ where $\lambda \in \mathbb{R}_{>0}$, $u \in \mathcal{U}(\mathcal{H}_\mathbb{C})$ is a unitary operator, $r = RP_S + Q_S$ and $v \in \mathfrak{v}$. In particular*

$$\mathcal{H}_\mathbb{C}^\bullet \simeq M_n^+ \times \exp(\mathfrak{v}) \times \mathbb{R}_{>0} \times \mathcal{U}(\mathcal{H}_\mathbb{C}).$$

The manifold Σ can be regarded as a universal model for the symmetric spaces of the noncompact type, namely

Theorem F. *For any finite dimensional real symmetric manifold M of the noncompact type (i.e., with no Euclidean de Rham factor, simply connected and with nonpositive sectional curvature), there is an embedding $M \hookrightarrow \Sigma$ which is a diffeomorphism between M and a closed geodesically convex submanifold of Σ. If we pull back the inner product on Σ to M, this inner product is a positive constant multiple of the inner product of M on each irreducible de Rham factor.*

The proof of the theorem is straightforward fixing an orthonormal basis of H (see Section 7.1) and recalling the well known result [13] that for any such space M there is an almost isometric embedding of M into $GL(\mathfrak{g})^+$, where $\mathfrak{g}$ is the Lie algebra of the Lie group $I_0(M)$ (the connected component of the identity of the group of isometries of M).

2. Background and definitions

Let $B(H)$ be the set of bounded operators acting on a complex, infinite dimensional and separable Hilbert space H, and let HS be the bilateral ideal of Hilbert–Schmidt operators of $B(H)$. Recall that HS is a Banach algebra (without unit) when given the norm $\|a\|_2 = \mathrm{tr}(a^*a)^{1/2}$ (see [26] for a detailed exposition on trace-class ideals). We will use HS^h to denote the closed subspace of self-adjoint Hilbert–Schmidt operators. In $B(H)$ we define

$$\mathcal{H}_\mathbb{C} = \{a + \lambda:\ a \in \mathsf{HS}, \lambda \in \mathbb{C}\},$$

the complex linear subalgebra consisting of Hilbert–Schmidt perturbations of scalar multiples of the identity (the closure of this algebra in the operator norm is the set of compact perturbations of scalar multiples of the identity). There is a natural Hilbert space structure for this subspace (where scalar operators are orthogonal to Hilbert–Schmidt operators) which is given by the inner product

$$\langle a + \lambda, b + \beta \rangle_2 = \mathrm{tr}(ab^*) + \lambda \bar{\beta}.$$

The algebra $\mathcal{H}_\mathbb{C}$ is complete with this norm. The model space that we are interested in is the real part of $\mathcal{H}_\mathbb{C}$,

$$\mathcal{H}_\mathbb{R} = \{a + \lambda:\ a^* = a,\ a \in \mathsf{HS},\ \lambda \in \mathbb{R}\},$$

which inherits the structure of (real) Banach space, and with the same inner product, becomes a real Hilbert space.

Remark 2.1. By virtue of trace properties, $\langle xy, y^*x^*\rangle_2 = \langle yx, x^*y^*\rangle_2$ for any $x, y \in \mathcal{H}_\mathbb{C}$, and also $\langle zx, yz\rangle_2 = \langle xz, zy\rangle_2$ for $x, y \in \mathcal{H}_\mathbb{C}$ and $z \in \mathcal{H}_\mathbb{R}$.

Let $\Sigma := \{A > 0\colon A \in \mathcal{H}_\mathbb{R}\}$ be the subset of positive invertible operators in $\mathcal{H}_\mathbb{R}$. It is clear that Σ is an open set of $\mathcal{H}_\mathbb{R}$ (for instance, using the lower semi continuity of the spectrum).

Remark 2.2. For $p \in \Sigma$, we identify $T_p\Sigma$ with $\mathcal{H}_\mathbb{R}$, and endow this manifold with a (real) Riemannian metric by means of the formula

$$\langle x, y\rangle_p := \langle p^{-1}x, yp^{-1}\rangle_2 = \langle xp^{-1}, p^{-1}y\rangle_2.$$

Throughout, let $\|x\|_p := \langle x, x\rangle_p^{1/2}$. Equivalently, $\|x\|_p = \|p^{-1/2}xp^{-1/2}\|_2$.

Lemma 2.3. *The covariant derivative in Σ (for the metric introduced in Remark 2.2) is given by*

$$\{\nabla_X Y\}_p = \{X(Y)\}_p - \frac{1}{2}(X_p p^{-1} Y_p + Y_p p^{-1} X_p). \tag{3}$$

Here $X(Y)$ denotes derivation of the vector field Y in the direction of X performed in the linear space $\mathcal{H}_\mathbb{R}$.

Proof. Note that ∇ is clearly symmetric and verifies all the formal identities of a connection; the proof that it is the Levi-Civita connection relays on the compatibility condition between the connection and the metric, $\frac{d}{dt}\langle X, Y\rangle_\gamma = \langle \nabla_{\dot\gamma} X, Y\rangle_\gamma + \langle X, \nabla_{\dot\gamma} Y\rangle_\gamma$ (see for instance [19, Chapter VIII, Theorem 4.1]). Here γ is a smooth curve in Σ and X, Y are tangent vector fields along γ. This identity is straightforward from the definitions and the properties of the trace. $\square$

Let $r^\alpha = e^{\alpha \ln(r)}$ (here $r \in \Sigma, \alpha \in \mathbb{R}$). The exponential is given by the usual series; note that any positive invertible operator has a real analytic logarithm, which is the inverse of the exponential in the Banach algebra. Note that $aba > 0$ whenever $a, b > 0$ and also $r^\alpha > 0$ whenever $r > 0$ and $\alpha \in \mathbb{R}$.

Euler's equation $\nabla_{\dot\gamma}\dot\gamma = 0$ for the covariant derivative introduced above reads $\ddot\gamma = \dot\gamma\gamma^{-1}\dot\gamma$, and it is not hard to see that the (unique) solution of this equation with $\gamma(0) = p, \gamma(1) = q$ is given by the smooth curve

$$\gamma_{pq}(t) = p^{\frac{1}{2}}(p^{-\frac{1}{2}}qp^{-\frac{1}{2}})^t p^{\frac{1}{2}}. \tag{4}$$

Remark 2.4. We will use $\mathrm{Exp}_p: T_p\Sigma \to \Sigma$ to denote the exponential map of Σ. Differentiating at $t = 0$ the curve above, we obtain $\dot\gamma_{pq}(0) = p^{1/2}\ln(p^{-1/2}qp^{-1/2})p^{1/2}$, hence

$$\mathrm{Exp}_p^{-1}(q) = p^{\frac{1}{2}}\ln(p^{-\frac{1}{2}}qp^{-\frac{1}{2}})p^{\frac{1}{2}} \quad \text{and} \quad \mathrm{Exp}_p(v) = p^{\frac{1}{2}}\exp(p^{-\frac{1}{2}}vp^{-\frac{1}{2}})p^{\frac{1}{2}}.$$

Note that by the construction above the map $\mathrm{Exp}_p: T_p\Sigma \to \Sigma$ is surjective (for given $q \in \Sigma$ take $v = p^{1/2}\ln(p^{-1/2}q \times p^{-1/2})p^{1/2}$, then $\mathrm{Exp}_p(v) = q$). Rearranging the exponential series we get the expressions $\mathrm{Exp}_p(v) = pe^{p^{-1}v} = e^{vp^{-1}}p$.

Lemma 2.5. *The metric in Σ is invariant under the action of the group of invertible elements: if g is an invertible operator in $\mathcal{H}_\mathbb{C}$, then $I_g(p) = gpg^*$ is an isometry of Σ.*

Proof. First note that for any $\psi \in H$ we have $\langle gpg^*\psi, \psi\rangle = \langle pg^*\psi, g^*\psi\rangle = \langle p\eta, \eta\rangle > 0$ assuming $p > 0$ and g invertible, so I_g maps Σ into itself. Also note that $d(I_g)_r(x) = gxg^*$ for any $x \in T_r\Sigma$, hence

$$\|gxg^*\|_{grg^*}^2 = \langle gxg^*(g^*)^{-1}r^{-1}g^{-1}, (g^*)^{-1}r^{-1}g^{-1}gxg^*\rangle_2 = \langle gxr^{-1}g^{-1}, (g^*)^{-1}r^{-1}xg^*\rangle_2$$
$$= \langle xr^{-1}, r^{-1}x\rangle_2 = \|x\|_r^2$$

where the third equality in the above equation follows from Remark 2.1. $\square$

3. Local and global structure

3.1. Curvature

We start showing that curvature in this manifold is a measure of noncommutativity, and then give a few definitions, which are necessary because of the infinite dimensional setting. Let $[\,,\,]$ stand for the usual commutator of operators, $[x, y] = xy - yx$.

Proposition 3.1. *The curvature tensor for the manifold Σ is given by*:

$$\mathfrak{R}_p(x, y)z = -\frac{1}{4}\, p\big[[p^{-1}x, p^{-1}y], p^{-1}z\big]. \tag{5}$$

Proof. This follows from the usual definition $\mathfrak{R}(x, y) = \nabla_x \nabla_y - \nabla_y \nabla_x - \nabla_{[x,y]}$. The formula for ∇ was given in Lemma 2.3. □

Definition 3.2. A Riemannian submanifold $M \subset \Sigma$ is *flat* at $p \in M$ if the sectional curvature vanishes for any 2-subspace of $T_p M$. The manifold M is *flat* if it is flat at any $p \in M$. The manifold M is *geodesic* at $p \in M$ if geodesics of the ambient space starting at p with initial velocity in $T_p M$ are also geodesics of M. The manifold M is a *totally geodesic* manifold if it is geodesic at any $p \in M$. Equivalently, M is totally geodesic if any geodesic of M is also a geodesic of Σ.

Proposition 3.3. *The manifold Σ has nonpositive sectional curvature.*

Proof. Let $x, y \in T_p \Sigma$. Let $\bar{x} = p^{-1/2}xp^{-1/2}$, $\bar{y} = p^{-1/2}yp^{-1/2}$. We may assume that x, y are orthonormal at p. A straightforward computation shows that

$$\mathcal{S}_p(x, y) = \langle \mathfrak{R}_p(x, y)y, x \rangle_p = -\frac{1}{4}\big\{ \langle \bar{x}\bar{y}^2, \bar{x} \rangle_2 - 2\langle \bar{y}\bar{x}\bar{y}, \bar{x} \rangle_2 + \langle \bar{y}^2\bar{x}, \bar{x} \rangle_2 \big\}.$$

Since $\bar{x}, \bar{y} \in \mathcal{H}_\mathbb{R}$, $\bar{x} = \lambda + a$ and $\bar{y} = \beta + b$ for $\lambda, \beta \in \mathbb{R}$ and $a, b \in \mathsf{HS}$. The equation reduces to

$$\mathcal{S}_p(x, y) = -\frac{1}{2}\big\{ \mathrm{tr}(a^2 b^2) - \mathrm{tr}\big((ab)^2\big) \big\}. \tag{6}$$

Note that $[z, w] = \mathrm{tr}(w^* z)$ is an inner product on HS, so we have the Cauchy–Schwarz inequality $\mathrm{tr}(w^* z) \leqslant \mathrm{tr}^{1/2}(w^* w)\, \mathrm{tr}^{1/2}(z^* z)$. Putting $w = ba$, $z = ab$, we obtain

$$\mathrm{tr}\big((ab)^2\big) = \mathrm{tr}(abab) = \mathrm{tr}\big((ba)^* ab\big) \leqslant \mathrm{tr}^{\frac{1}{2}}(abba)\, \mathrm{tr}^{\frac{1}{2}}(baab) = \mathrm{tr}(a^2 b^2). \qquad \square$$

Proposition 3.4. *Let $M \subset \Sigma$ be a submanifold. Assume that M is flat and geodesic at $p \in M$. If $x, y \in T_p M$, then $p^{-1/2}xp^{-1/2}$ commutes with $p^{-1/2}yp^{-1/2}$.*

Proof. Since M is geodesic at p, the curvature tensor is the restriction of the curvature tensor of Σ, so in Eq. (6) above the right-hand term must be zero if M is flat at p. But the Cauchy–Schwarz inequality is an equality only if the vectors are linearly dependent; in the notation of the previous theorem, we have $ab = z = \alpha w = \alpha ba$ for some $\alpha \in \mathbb{R}$; replacing this in the above equation we obtain $\alpha = 1$, namely $ab = ba$. Recalling the definitions for a and b we obtain the assertion.

3.2. Convexity of Jacobi fields

Let $J(t)$ be a Jacobi field along a geodesic γ of Σ, i.e. J is a solution of the differential equation

$$D_t^2 J + \mathfrak{R}_\gamma(J, \dot{\gamma})\dot{\gamma} = 0 \tag{7}$$

where $D_t = \nabla_{\dot\gamma}$ is the covariant derivative along γ. We may assume that $J(t)$ is nonvanishing, hence

$$\|J\|_\gamma^3 \frac{d^2}{dt^2} \langle J, J\rangle_\gamma^{\frac{1}{2}} = -\langle D_t J, J\rangle_\gamma^2 + \langle J, J\rangle_\gamma \langle D_t J, D_t J\rangle_\gamma - \langle J, J\rangle_\gamma \langle \mathfrak{R}_\gamma(J, \dot\gamma)\dot\gamma, J\rangle_\gamma.$$

The third term is clearly positive and the first two terms add up to a nonnegative number by the Cauchy–Schwarz inequality: $\langle D_t J, J\rangle_\gamma^2 \leqslant \langle D_t J, D_t J\rangle_\gamma \langle J, J\rangle_\gamma$. In other words, the smooth function $t \mapsto \langle J, J\rangle_\gamma^{1/2} = \|J\|_\gamma$ is convex, exactly as in the finite dimensional setting.

3.3. The exponential map

We present two theorems that, in this infinite dimensional setting, stem from McAlpin's PhD thesis (for a proof see [21] or Theorem 3.7 of Chapter IX in [19]). First, if one identifies the Riemannian exponential with a suitable Jacobi lift, one obtains

Theorem 3.5. *The map* $\operatorname{Exp}_p : T_p\Sigma \to \Sigma$ *has an expansive differential*:

$$\left\| d(\operatorname{Exp}_p)_v(w) \right\|_{\operatorname{Exp}_p(v)} \geqslant \|w\|_p.$$

This result implies that the differential of the exponential map is injective and has closed range. Playing with the Hilbert structure of the tangent bundle and using the well-known identity for operators $\operatorname{Ker}(A)^\perp = \overline{\operatorname{Ran}(A^*)}$, it can be proved that this map is surjective, moreover

Corollary 3.6. *The differential of the Riemannian exponential* $d(\operatorname{Exp}_p)_v : T_p\Sigma \to T_{\operatorname{Exp}_p(v)}\Sigma$ *is a linear isomorphism for any* $v \in T_p\Sigma$. *Hence,* $\operatorname{Exp}_p : T_p M \to \Sigma$ *is a* C^ω-*diffeomorphism.*

The last assertion is due to the fact that the map $\operatorname{Exp}_p : T_p\Sigma \to \Sigma$ is a bijection (see Remark 2.4 above).

3.4. The shortest path and the geodesic distance

The following inequality is the key to the proof of the fact that geodesics are minimizing. It was proved by R. Bhatia [6] for matrices, and his proof can be translated almost verbatim to the context of operator algebras with a trace, see [2]. However since the Riemannian metric in Σ is complete, the inequality can be easily deduced from the fact that the norm of a Jacobi field is a convex map (in Theorem 3.5 put $p = 1$, $v = x$ and $w = y$):

Corollary 3.7. *If* $d\exp_x$ *denotes the differential at* x *of the usual exponential map, then for any* $x, y \in \mathcal{H}_\mathbb{R}$

$$\left\| d\exp_x(y) \right\|_{e^x} = \left\| e^{-\frac{x}{2}} d\exp_x(y) e^{-\frac{x}{2}} \right\|_2 \geqslant \|y\|_2.$$

As usual, one measures length of curves in Σ using the norms in each tangent space,

$$L(\alpha) = \int_0^1 \|\dot\alpha(t)\|_{\alpha(t)} \, dt. \tag{8}$$

We define the distance between two points $p, q \in \Sigma$ as the infimum of the lengths of piecewise smooth curves in Σ joining p to q,

$$\operatorname{dist}(p, q) = \inf \left\{ L(\alpha) : \alpha \subset \Sigma, \ \alpha(0) = p, \ \alpha(1) = q \right\}.$$

Recall (Remark 2.4 and the paragraph above it) that for any pair of elements $p, q \in \Sigma$, we have the smooth curve $\gamma_{pq} \subset \Sigma$, $\gamma_{pq}(t) = p^{1/2}(p^{-1/2}qp^{-1/2})^t p^{1/2}$ joining p to q, which is the unique solution of Euler's equation in Σ. Computing the derivative, we get

$$\left\| \dot\gamma_{pq}(t) \right\|_{\gamma_{pq}(t)} \equiv \left\| \ln(p^{\frac{1}{2}} q^{-1} p^{\frac{1}{2}}) \right\|_2 = L(\gamma_{pq}).$$

The minimality of these (unique) geodesics joining two points can be deduced from general considerations [14], we present here a direct proof.

Theorem 3.8. *Let* $p, q \in \Sigma$. *Then the geodesic* γ_{pq} *is the shortest curve joining* p *and* q *in* Σ, *if the length of curves is measured with the metric defined above* (8).

Proof. Let α be a smooth curve in Σ with $\alpha(0) = p$ and $\alpha(1) = q$. We must compare the length of α with the length of γ_{pq}. Since the invertible group acts isometrically, it preserves the lengths of curves. Thus we may act with $p^{-1/2}$, and suppose that both curves start at 1, or equivalently that $p = 1$. Therefore $\gamma_{1q}(t) := \gamma(t) = e^{tx}$, with $x = \ln q$. The length of γ is then $\|x\|_2$. The proof follows easily from the inequality of Corollary 3.7. Indeed, since α is a smooth curve in Σ, it can be written as $\alpha(t) = e^{\beta(t)}$, with $\beta = \ln \alpha$. Then β is a smooth curve of self-adjoint operators with $\beta(0) = 0$ and $\beta(1) = x$. Moreover,

$$L(\gamma) = \|x\|_2 = \|x - 0\|_2 = \left\| \int_0^1 \dot\beta(t)\, dt \right\|_2 \leqslant \int_0^1 \|\dot\beta(t)\|_2\, dt.$$

On the other hand, by the mentioned inequality,

$$\|\dot\beta(t)\|_2 \leqslant \|e^{-\frac{\beta(t)}{2}} d\exp_{\beta(t)}(\dot\beta(t)) e^{-\frac{\beta(t)}{2}}\|_2 = \|d\exp_{\beta(t)}(\dot\beta(t))\|_{e^{\beta(t)}} = \|\dot\alpha(t)\|_{\alpha(t)}. \qquad \square$$

Remark 3.9. The geodesic distance induced by the metric is given by

$$\mathrm{dist}(p, q) = \|\ln(p^{\frac{1}{2}} q^{-1} p^{\frac{1}{2}})\|_2.$$

Hence the unique geodesic joining p to q is also the shortest path joining p to q. This means that (Σ, dist) is a (not locally compact) *geodesic length space* in the sense of Aleksandrov and Gromov [5]. These curves look formally equal to the geodesics between positive definite $n \times n$ matrices, when this space is regarded as a symmetric space.

Corollary 3.10. *If* γ, δ *are geodesics, the map* $f : \mathbb{R} \to \mathbb{R}$, $t \mapsto \mathrm{dist}(\gamma(t), \delta(t))$ *is convex.*

Proof. The distance between the points $\gamma(t)$ and $\delta(t)$ is given by the geodesic $\alpha_t(s)$, which is obtained as the s variable ranges in a geodesic square $h(s, t)$ with vertices $\{\gamma(t_0), \delta(t_0), \gamma(t_1), \delta(t_1)\}$ (the starting and ending points of γ and δ). Taking the partial derivative along the direction of s gives a Jacobi field $J(s, t)$ along the geodesic $\beta_s(t) = h(s, t)$ and it also gives the speed of α_t. Hence

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

This equation states that $f(t)$ can be written as the limit of a convex combination of convex functions $u_i(t) = \|J(s_i, t)\|_{h(s_i,t)}$, so f must be convex itself. $\square$

In a recent paper (Corollary 8.7 of [20]), the authors prove this property of convexity of the geodesic distance in a general setting concerning nonpositively curved symmetric spaces given by a quotient of Banach–Lie groups.

Lemma 3.11. *For any* $x, y \in \mathcal{H}_{\mathbb{R}}$ *we have*

$$\mathrm{dist}(e^x, e^y) = \|\ln(e^{x/2} e^{-y} e^{x/2})\|_2 \geqslant \|x - y\|_2. \tag{9}$$

Proof. Take $\gamma(t) = e^{tx}$, $\delta(t) = e^{ty}$ and f as in the previous corollary; we may assume that $x, y \in \mathsf{HS}^h$. Note that $f(0) = 0$, hence $f(t)/t \leqslant f(1)$ for any $0 < t \leqslant 1$; hence $\lim_{t\to 0^+} f(t)/t \leqslant f(1)$. Now

$$f(t)/t = \frac{1}{t} \|\ln(e^{tx/2} e^{-ty} e^{tx/2})\|_2 = \mathrm{tr}\left(\left[\frac{1}{t}\ln(e^{tx/2} e^{-ty} e^{tx/2})\right]^2\right)^{\frac{1}{2}},$$

and

$$\lim_{t \to 0^+} \frac{1}{t} \ln(e^{tx/2} e^{-ty} e^{tx/2}) = \frac{d}{dt} \bigg|_{t=0} \ln(e^{tx/2} e^{-ty} e^{tx/2}) = d \ln_1(x - y) = x - y. \qquad \square$$

Corollary 3.12. *The inner angles of any geodesic triangle in Σ add up to at most π.*

Proof. Using the invariance of the metric for the action of the group of invertible operators, and squaring both sides of inequality (9) in Lemma 3.11, we obtain the Hyperbolic Cosine Law:

$$l_i^2 \geqslant l_{i+1}^2 + l_{i-1}^2 - 2l_{i+1}l_{i-1}\cos(\alpha_i). \tag{10}$$

Here l_i ($i = 1, 2, 3$) are the sides of any geodesic triangle and α_i is the angle opposite to l_i. These inequalities put together show that one can construct a comparison Euclidean triangle in the affine plane with sides l_i. For this triangle with angles β_i (opposite to the side l_i) we have $l_i^2 = l_{i+1}^2 + l_{i-1}^2 - 2l_{i+1}l_{i-1}\cos(\beta_i)$. This equation together with inequality (10) imply that the angle β_i is bigger than α_i for $i = 1, 2, 3$. Adding the three angles we have $\alpha_1 + \alpha_2 + \alpha_3 \leqslant \beta_1 + \beta_2 + \beta_3 = \pi$. $\square$

Proposition 3.13. *The metric space (Σ, d) is complete with the distance induced by the minimizing geodesics.*

Proof. Consider a Cauchy sequence $\{p_n\} \subset \Sigma$. Again by virtue of inequality (9) of Lemma 3.11, $x_n = \ln(p_n)$ is a Cauchy sequence in $\mathcal{H}_\mathbb{R}$. Since Hilbert–Schmidt operators are complete with the trace norm, there is a vector $x \in \mathcal{H}_\mathbb{R}$ such that $x_n \to x$ in the trace norm. Since the inverse map, the exponential map, the product and the logarithm are all analytic maps with respect to the trace norm, $\mathrm{dist}(p_n, e^x) = \|\ln(e^{x/2}e^{-x_n}e^{x/2})\|_2 \to 0$ when $n \to \infty$. $\square$

4. Geodesically convex submanifolds

Definition 4.1. A set $M \subset \Sigma$ is *geodesically convex* if for any two given points $p, q \in M$, the unique geodesic of Σ joining p to q lies entirely in M. A Riemannian submanifold $M \subset \Sigma$ is complete at $p \in M$ if Exp_p^M is defined in the whole tangent space and maps onto M. The manifold M is *complete* if it is complete at any point.

Remark 4.2. The manifold Σ is complete; moreover, Exp_p is a C^ω (analytic) isomorphism of $\mathcal{H}_\mathbb{R}$ with Σ for each $p \in \Sigma$. Other notions of completeness are touchy because, as C.J. Atkin shows in [3] and [4], the Hopf–Rinow Theorem does not necessarily hold in infinite dimensional Banach manifolds.

These previous notions are strongly related; it is not hard to see that for any Riemannian submanifold M of Σ, M is geodesically convex if and only if M is complete and totally geodesic. On the other hand, it should be clear from the definitions that whenever M is a convex submanifold of Σ, M is nonpositively curved.

4.1. An intrinsic characterization of convexity

From now on the term *convex* stands for the longer *geodesically convex*. As before $[\ ,\ ]$ denotes the usual commutator of operators in $B(H)$. To deal with convex sets the following definition will be useful; assume $\mathfrak{m} \subset \mathcal{H}_\mathbb{R}$ is a real linear space.

Definition 4.3. We say that $\mathfrak{m}$ is a *Lie triple system* if $[[a, b], c] \in \mathfrak{m}$ for any $a, b, c \in \mathfrak{m}$. Equivalently, $[x, [x, y]] \in \mathfrak{m}$ whenever $x, y \in \mathfrak{m}$.

Note that whenever a, b, c are self-adjoint operators, $d = [a, [b, c]]$ is also a self-adjoint operator. So, for any involutive Lie subalgebra of operators $\mathfrak{a} \subset \mathcal{H}_\mathbb{C}$ (in particular: for any associative Banach subalgebra), $\mathfrak{m} = \mathfrak{Re}(\mathfrak{a})$ is a Lie triple system in $\mathcal{H}_\mathbb{R}$.

Assume $M \subset \Sigma$ is a submanifold such that $1 \in M$, and M is geodesic at $p = 1$. Then $T_1 M$ is a Lie triple system, because the curvature tensor at $p = 1$ is the restriction to $T_1 M$ of the curvature tensor of Σ, and $\mathfrak{R}_1(x, y)z =$

$-\frac{1}{4}[[x, y], z]$. In particular, if M is geodesically convex, $T_1 M$ must be a Lie triple system. This weak condition on the tangent space turns out to be strong enough to obtain convexity:

Theorem 4.4. *(See Mostow–de la Harpe [16,22].) Assume* $\mathfrak{m} \subset \mathcal{H}_\mathbb{R}$ *is a closed subspace, put* $M = \exp(\mathfrak{m}) \subset \Sigma$ *with the induced topology and Riemannian metric. Assume further that* $\mathfrak{m}$ *is a Lie triple system. Then for any* $p, q \in M$ *it holds true that* $qpq \in M$.

Proof. As P. de la Harpe pointed out, the proof of G.D. Mostow for matrices in [22] can be translated to Hilbert–Schmidt operators without any modification: we give a sketch of the proof here. Assume $p = e^x$, $q = e^y$, and consider the curve $e^{\alpha(t)} = e^{ty} e^x e^{ty}$. Then it can be proved that $\dot{\alpha}(t) = G(\alpha(t))$ with G a Lipschitz map that sends $\mathfrak{m}$ into $\mathfrak{m}$ (this is nontrivial). Since $\alpha(0) = x \in \mathfrak{m}$ and G is a Lipschitz map by the uniqueness of the solutions of ordinary differential equations we have $\alpha \subset \mathfrak{m}$. Hence $e^{\alpha(1)} = qpq \in M$ and the claim follows. □

Corollary 4.5. *Assume* $M = \exp(\mathfrak{m}) \subset \Sigma$, *and* $\mathfrak{m}$ *is as in the above theorem. Then* M *is a closed convex submanifold.*

Proof. Take $p, q \in M$. Then $p = e^x$, $q = e^y$ with $x, y \in \mathfrak{m}$. If we put $r = e^{-x/2} e^y e^{-x/2}$, then $r \in M$ because $e^{-x/2}$ and e^y are in M. Moreover, $z = \ln(r) \in \mathfrak{m}$. But the unique geodesic of Σ joining p to q is $\gamma(t) = e^{x/2} e^{tz} e^{x/2}$, hence $\gamma \subset M$. □

Corollary 4.6. *Assume* $\mathfrak{m} \subset \mathcal{H}_\mathbb{R}$ *is a closed, commutative associative Banach subalgebra of* $\mathcal{H}_\mathbb{C}$. *Then the manifold* $M = \exp(\mathfrak{m}) \subset \Sigma$ *is a closed, convex and flat Riemannian submanifold. Moreover,* M *is an open subset of* $\mathfrak{m}$ *and an abelian Banach–Lie group.*

Proof. The first assertion follows from the fact that $\mathfrak{m}$ is a Lie triple system. Curvature is given by commutators, hence M is flat. Since $\mathfrak{m}$ is a closed subalgebra, $e^x = \sum \frac{x^n}{n!} \in \mathfrak{m}$ for any $x \in \mathfrak{m}$, so $M \subset \mathfrak{m}$. That M is open follows from the fact that exp is a C^ω isomorphism (Corollary 3.6).

If M is flat and geodesic at $p = 1$, $T_1 M = \mathfrak{m}$ is abelian (by Proposition 3.4), therefore

Corollary 4.7. *Assume* $M = \exp(\mathfrak{m})$ *is closed and flat. If* M *is geodesic at* $p = 1$, *then* M *is a convex submanifold. Moreover,* M *is an abelian Banach–Lie group and an open subset of* $\mathfrak{m}$.

We adopt the usual definition of a symmetric space [15]:

Definition 4.8. A Riemann–Hilbert manifold M is called a *globally symmetric space* if each point $p \in M$ is an isolated fixed point of an involutive isometry $s_p : M \to M$. The map s_p is called the *geodesic symmetry*.

Theorem 4.9. *Assume* $M = \exp(\mathfrak{m})$ *is closed and convex. Then* M *is a symmetric space; the geodesic symmetry at* $p \in M$ *is given by* $s_p(q) = pq^{-1}p$ *for any* $q \in M$. *In particular,* Σ *is a symmetric space.*

Proof. Observe that, for $p = e^x$, $q = e^y$, $s_p(q) = e^x e^{-y} e^x$; this shows that s_p maps M into M. To prove that s_p is an isometry, for any vector $v \in \mathfrak{m}$ consider the geodesic α_v of M such that $\alpha(0) = q$ and $\dot{\alpha}(0) = v$. Then $\alpha(t) = q e^{t q^{-1} v}$ and

$$d(s_p)_q(v) = \frac{d}{dt}\bigg|_{t=0} (s_p \circ \alpha_v) = -pq^{-1} v q^{-1} p.$$

Since M has the induced metric, $\|pq^{-1} v q^{-1} p\|^2_{pq^{-1}p} = \|v\|^2_q$ by Lemma 2.5 (with $g = pq^{-1}$). In particular, $d_p s_p = -id$, so p is an isolated fixed point of s_p for any $p \in M$. □

Theorem 4.4 and its corollaries imply that Σ (as any symmetric space) contains plenty of convex sets; in particular

Remark 4.10. We can embed isometrically any k-dimensional plane in Σ as a convex closed submanifold: take an orthonormal set of k commuting operators (for instance, fix an orthonormal basis $\{e_i\}_{i\in\mathbb{M}}$ of H and take $p_i = e_i \otimes e_i$, $i = 1, \ldots, k$), and consider the exponential of the linear span of this set. In the language of symmetric spaces, we are saying that $\operatorname{rank}(\Sigma) = +\infty$.

Let $I(M)$ be the group of isometries of a submanifold M.

Theorem 4.11. *If the submanifold* $M = \exp(\mathfrak{m})$ *is closed and convex, then* $I(M)$ *acts transitively on* M.

Proof. Take $p = \mathrm{e}^x$, $q = \mathrm{e}^y$ two points in M, $v = p \ln(p^{-1}q)$ and $\gamma(t) = p\mathrm{e}^{t\, p^{-1}v}$ the geodesic joining p to q. Note that $q = \gamma(1) = p\mathrm{e}^{p^{-1}v} = \mathrm{e}^{vp^{-1}}p$. Consider the curve of isometries $\varphi_t = s_{\gamma(t/2)} \circ s_p$. Then

$$\varphi_1(p) = \mathrm{e}^{\frac{1}{2}v\mathrm{e}^{-x}}\mathrm{e}^x\mathrm{e}^{-x}\,\mathrm{e}^{\frac{1}{2}v\mathrm{e}^{-x}}\mathrm{e}^x = \mathrm{e}^{v\mathrm{e}^{-x}}\mathrm{e}^x = q. \qquad \square$$

Remark 4.12. Assume $M \subset \Sigma$ is closed and convex, and $1 \in M$. Let $I(M)$ be the group of isometries of M. Then, since any isometry φ is uniquely determined by its value at $1 \in M$ and its differential $d\varphi_1$, the set $I(M)$ can be naturally embedded in a Banach space: take $\varphi \in I(M)$ and consider

$$\overline{\varphi}\,(q) = \varphi(1)^{-\frac{1}{2}}\,\varphi(q)\,\varphi(1)^{-\frac{1}{2}}.$$

Note that $d\overline{\varphi}_1$ is a unitary operator of $T_1 M = \mathfrak{m}$ (with the natural Hilbert-space structure), so there is an inclusion $J: I(M) \hookrightarrow M \times \mathcal{U}(\mathrm{B}(\mathfrak{m}))$ given by the map $\varphi \mapsto (\varphi(1), d\overline{\varphi}_1)$. On the other hand, for a given pair $(p, u) \in M \times \mathcal{U}(\mathrm{B}(\mathfrak{m}))$, put $\varphi(\mathrm{e}^x) = p^{1/2} \exp(u(x)) p^{1/2} (x \in \mathfrak{m})$. It is not hard to see that φ is an isometry of M which maps 1 to p, such that $d\overline{\varphi}_1 = u$. Hence we may identify $I(M) \simeq M \times \mathcal{U}(\mathrm{B}(\mathfrak{m}))$.

Remark 4.13. If $M = \exp(\mathfrak{m})$ is closed and convex, it is geodesic at any $p = \mathrm{e}^x \in M$, so

$$T_p M = \mathrm{Exp}_p^{-1}(M) = \left\{ p^{\frac{1}{2}}\ln(p^{-\frac{1}{2}} q\, p^{-\frac{1}{2}})p^{\frac{1}{2}}\colon q \in M \right\}$$

(see Remark 2.4). Since $p^{1/2} = \mathrm{e}^{x/2} \in \mathfrak{m}$, using Theorem 4.4 we obtain the identification $T_p M = p^{1/2}(T_1 M)p^{1/2} = p^{1/2} \mathfrak{m}\, p^{1/2}$. It also follows easily that an operator $v \in \mathcal{H}_\mathbb{R}$ is orthogonal to M at p (that is, $v \in T_p M^\perp$) if and only if

$$\langle p^{-\frac{1}{2}} z\, p^{-\frac{1}{2}}\,,\, v\rangle_2 = \langle p^{-\frac{1}{2}} v\, p^{-\frac{1}{2}}, z\rangle_2 = 0 \quad \text{for any } z \in \mathfrak{m}.$$

In particular, $T_1 M^\perp = \mathfrak{m}^\perp = \{v \in \mathcal{H}_\mathbb{R}\colon \langle v, z\rangle_2 = 0 \text{ for any } z \in \mathfrak{m}\}$. Note that, when $\mathfrak{m}$ is a closed commutative *associative subalgebra* of operators, $y \mapsto p^{1/2} y p^{1/2}$ is a linear isomorphism of $\mathfrak{m}$; in this case $T_p M = \mathfrak{m} = T_1 M$ for any $p \in M$. This last assertion also follows easily from Corollary 4.6, and clearly $T_p M^\perp = T_1 M^\perp = \mathfrak{m}^\perp$ in this case.

Remark 4.14. Assume $M \subset \Sigma$ is a convex submanifold. If the curve γ is the geodesic joining p to q, then the isometry $\varphi_t = s_{\gamma(t/2)} \circ s_p$ translates along γ, namely

$$\varphi_t\big(\gamma(s)\big) = p\mathrm{e}^{\frac{t}{2}p^{-1}v}p^{-1}p\mathrm{e}^{sp^{-1}v}p^{-1}p\mathrm{e}^{\frac{t}{2}p^{-1}v} = p\mathrm{e}^{\frac{t}{2}p^{-1}v}\mathrm{e}^{sp^{-1}v}\mathrm{e}^{\frac{t}{2}p^{-1}v} = p\mathrm{e}^{(s+t)p^{-1}v} = \gamma(s+t).$$

In particular, $\varphi_1(p) = q$. Now take any tangent vector $w \in T_{\gamma(s)}M$, and let

$$w(t) = (d\varphi_t)_{\gamma(s)}(w) = \mathrm{e}^{\frac{t}{2}vp^{-1}}w\mathrm{e}^{\frac{t}{2}p^{-1}v}.$$

It follows from a straightforward computation using Eq. (3) of Section 2 that $w(t)$ is the parallel translation of w from $\gamma(s)$ to $\gamma(s+t)$; namely $\nabla_{\dot\gamma} w \equiv 0$. We conclude that the linear map $(d\varphi_t)_{\gamma(s)}\colon T_{\gamma(s)}M \to T_{\gamma(s+t)}M$ gives parallel translation along γ, i.e. $(d\varphi_t)_{\gamma(s)} = P_s^{t+s}(\gamma)$. In particular, since $q = \gamma(1) = p^{1/2}\mathrm{e}^{p^{-1/2}vp^{1/2}}p^{1/2}$, the map

$$P_p^q\colon w \mapsto p^{\frac{1}{2}}(p^{-\frac{1}{2}}qp^{-\frac{1}{2}})^{\frac{1}{2}}p^{-\frac{1}{2}}wp^{-\frac{1}{2}}(p^{-\frac{1}{2}}qp^{-\frac{1}{2}})^{\frac{1}{2}}p^{\frac{1}{2}}$$

gives parallel translation from $T_p M$ to $T_q M$. See also Theorem 4.18.

4.1.1. Examples of convex sets

1. For any subspace $\mathfrak{s} \subset \mathcal{H}_{\mathbb{R}}$, $\mathfrak{m}_{\mathfrak{s}} = \{x \in \mathcal{H}_{\mathbb{R}}: [x, y] = 0 \ \forall y \in \mathfrak{s}\}$ is a Lie triple system.
2. In particular, for any $y \in \mathcal{H}_{\mathbb{R}}$, $\mathfrak{m}_y = \{x \in \mathcal{H}_{\mathbb{R}}: [x, y] = 0\}$ is a Lie triple system.
3. The family of operators in $\mathcal{H}_{\mathbb{R}}$ which act as endomorphisms of a closed subspace $S \subset H$ form a Lie triple system in $\mathcal{H}_{\mathbb{R}}$.
4. Any norm closed commutative associative subalgebra of $\mathcal{H}_{\mathbb{R}}$, closed under the usual involution of operators, is a Lie triple system. In particular
    (a) The diagonal operators (see Section 6). This is a maximal abelian closed subspace of $\mathcal{H}_{\mathbb{R}}$, hence the manifold Δ (which is the exponential of this set) is a maximal flat submanifold of Σ.
    (b) The scalar manifold $\Lambda = \{\lambda 1: \lambda \in \mathbb{R}_{>0}\} \subset \Sigma$ is the exponential of the Lie triple system $\mathbb{R} 1 \subset \mathcal{H}_{\mathbb{R}}$.
    (c) For fixed $a \in \mathsf{HS}^h$, the real part of the closed algebra generated by a, which is the closure in the 2-norm of the set of polynomials in a, is a Lie triple system.
5. The real part of any Banach–Lie subalgebra of $\mathcal{H}_{\mathbb{C}}$ is a Lie triple system (in particular: the real part of any associative Banach subalgebra).

4.2. Convex manifolds as homogeneous manifolds

The results of this section are related to those of Sections 3 and 7 of Chapter IV in [15]. See also Theorem 5.5 in [23] for a proof of the existence of smooth polar decompositions in the (broader) Banach–Finsler context.

Definition 4.15. Let $\mathcal{H}_{\mathbb{C}}^{\bullet}$ be the group of invertible elements in $\mathcal{H}_{\mathbb{C}}$. This group has a natural structure of manifold as an open set of the associative Banach algebra $\mathcal{H}_{\mathbb{C}}$; it is a Banach–Lie group with Banach–Lie algebra $\mathcal{H}_{\mathbb{C}}$.

Let $\mathcal{U}(\mathcal{H}_{\mathbb{C}})$ stand for the unitary elements of the involutive Banach algebra $\mathcal{H}_{\mathbb{C}}$, namely the set of $u \in \mathcal{H}_{\mathbb{C}}^{\bullet}$ such that $u^* = u^{-1}$. It is a real Banach–Lie subgroup of $\mathcal{H}_{\mathbb{C}}^{\bullet}$ with Lie algebra $i\mathcal{H}_{\mathbb{R}}$.

Let G be a connected abstract subgroup of $\mathcal{H}_{\mathbb{C}}^{\bullet}$. We say that G is a self-adjoint subgroup of $\mathcal{H}_{\mathbb{C}}^{\bullet}$ if $g^* \in G$ whenever $g \in G$ (for short, $G^* = G$). Note that a connected Banach–Lie group G is self-adjoint if and only if $\mathfrak{g}^* = \mathfrak{g}$, where $\mathfrak{g}$ denotes the Banach–Lie algebra of G.

If $\mathfrak{a} \subset \mathcal{H}_{\mathbb{C}}$ is a linear space over $\mathbb{R}$, let $[\mathfrak{a}, \mathfrak{a}] = \overline{\text{span}_{\mathbb{R}}\{[a, b]: a, b \in \mathfrak{a}\}}$, where the bar denotes closure in the norm of the Banach algebra $\mathcal{H}_{\mathbb{C}}$.

If $A \subset \mathcal{H}_{\mathbb{C}}^{\bullet}$ is a set, $\langle A \rangle$ will denote the abstract subgroup generated by A (the group whose elements are the inverses and the finite products of elements in A).

Let $|x| = (xx^*)^{1/2} = \exp(\frac{1}{2}\ln(xx^*))$ for $x \in \mathcal{H}_{\mathbb{C}}$. Since $\mathcal{H}_{\mathbb{C}}$ is an involutive Banach algebra, $|x| \in \Sigma \subset \mathcal{H}_{\mathbb{C}}^{\bullet}$ if $x \in \mathcal{H}_{\mathbb{C}}$.

Remark 4.16. The group $\mathcal{H}_{\mathbb{C}}^{\bullet}$, having the homotopy type of the inductive limit of the groups $GL(n, \mathbb{C})$ (see [16, Section II.6]) is connected; moreover, there is a homotopy class equivalence

$$\mathcal{H}_{\mathbb{C}}^{\bullet} \simeq S^1 \times S^1 \times SU(\infty).$$

Here $SU(\infty)$ stands for the inductive limit of the groups $SU(n, \mathbb{C})$.

Proposition 4.17. *Let $\mathfrak{g} \subset \mathcal{H}_{\mathbb{C}}$ be a closed real Banach–Lie subalgebra. Then $G = \langle \exp(\mathfrak{g}) \rangle$ admits a topology and a smooth structure such that G is a connected real Banach–Lie group and $\mathfrak{g} = T_1 G$ is the Banach–Lie algebra of G. The inclusion $G \hookrightarrow \mathcal{H}_{\mathbb{C}}^{\bullet}$ is a smooth immersion and the exponential map of G is given by the usual exponential of $\mathcal{H}_{\mathbb{C}}$. The topology on G might be strictly finer than the topology of $\mathcal{H}_{\mathbb{C}}^{\bullet}$.*

Proof. Since $\mathcal{H}_{\mathbb{C}}$ is a Hilbert space, the Banach–Lie subalgebra admits a supplement. By Theorem 5.4 of Chapter VI in [19], there exists an integral manifold $H \xrightarrow{j} \mathcal{H}_{\mathbb{C}}^{\bullet}$ for the subbundle $\{g\mathfrak{g}\}_{g \in \mathcal{H}_{\mathbb{C}}^{\bullet}}$. The manifold H is connected, and a Banach–Lie group with $dj_1(T_1 H) = \mathfrak{g}$. Since j is a smooth homomorphism of Banach–Lie groups, we have $j \circ \text{Exp}^H = \exp \circ dj_1$. The other assertions follow from this identity because $G = \langle \exp(\mathfrak{g}) \rangle = \langle j \circ \text{Exp}^H(T_1 H) \rangle = j(H)$. $\quad \square$

Theorem 4.18. *Let $G = \langle \exp(\mathfrak{g}) \rangle \subset \mathcal{H}_{\mathbb{C}}^{\bullet}$ be a connected self-adjoint Banach–Lie group with Banach–Lie algebra $\mathfrak{g} \subset \mathcal{H}_{\mathbb{C}}$. Let P be the analytic map $g \mapsto gg^*$, $P : G \to G$. Let $\mathfrak{k} = \ker(dP_1)$, $\mathfrak{m} = \mathrm{Ran}(dP_1)$. Let $M_G = \exp(\mathfrak{m})$, $K = G \cap \mathcal{U}(\mathcal{H}_{\mathbb{C}}) = P^{-1}(1)$. Then*

1. *The set $\mathfrak{m}$ is a closed Lie triple system in $\mathcal{H}_{\mathbb{R}}$. We have $[\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{k}$, $[\mathfrak{m}, \mathfrak{k}] \subset \mathfrak{m}$, $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$ and $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{k}$. In particular, $\mathfrak{k}$ is a Banach–Lie subalgebra of $\mathfrak{g}$ (and of $i\mathcal{H}_{\mathbb{R}}$ also).*
2. *$P(G) = M_G$, and M_G is a geodesically convex submanifold of Σ.*
3. *For any $g = |g| u_g \in G$ (polar decomposition), we have $|g| \in M_G$ and $u_g \in K$.*
4. *Let $g \in G$, $p \in M_G$, $I_g(p) = gpg^*$. Then $I_g \in I(M_G)$. If $g = p^{1/2}(p^{-1/2}qp^{-1/2})^{1/2}p^{-1/2} \in G$, then $I_g(p) = q$, namely G acts isometrically and transitively on M_G.*
5. *Let $u \in K$ and $x \in \mathfrak{m}$ (resp. $\mathfrak{m}^{\perp}$). Then $I_u(x) = uxu^* \in \mathfrak{m}$ (resp. $\mathfrak{m}^{\perp}$). If $p, q \in M_G$ then I_p maps $T_q M_G$ (resp. $T_q M_G{}^{\perp}$) isometrically onto $T_{I_p(q)} M_G$ (resp. $T_{I_p(q)} M_G{}^{\perp}$).*
6. *The group K is a Banach–Lie subgroup of G with Lie algebra $\mathfrak{k}$.*
7. *$G \simeq M_G \times K$ as Hilbert manifolds. In particular K is connected and $G/K \simeq M_G$.*

Proof. 1. Note that $dP_1(x) = x + x^*$, hence $\mathfrak{k} = \{x \in \mathfrak{g}: x^* = -x\}$ which is certainly a closed Lie algebra. Note also that $\mathfrak{m} = \{x \in \mathfrak{g}: x^* = x\}$ is a Lie triple system; it is closed because $x \mapsto x^*$ is an isometric automorphism of $\mathcal{H}_{\mathbb{C}}$. Since $[x, y] = xy - yx$ is self-adjoint whenever x is self-adjoint and y is skew-adjoint, the other assertions are clear.

2. Clearly $P(G) \supseteq \exp(\mathfrak{m})$ because $e^x = P(e^{x/2})$. On the other hand, since $\mathfrak{g}$ splits, there exist neighborhoods of zero $U_{\mathfrak{m}} \subset \mathfrak{m}$ and $U_{\mathfrak{k}} \subset \mathfrak{k}$ such that the map $x_{\mathfrak{m}} + y_{\mathfrak{k}} \mapsto e^{x_{\mathfrak{m}}} e^{y_{\mathfrak{k}}}$ is an isomorphism from $U_{\mathfrak{m}} \oplus U_{\mathfrak{k}}$ onto an open neighborhood V of $1 \in G$. Then $\langle V \rangle$ is open (and closed) in G and so is all of G. Hence, for any $g \in G$, $g = (e^{x_1} e^{y_1})^{\alpha_1} \cdots (e^{x_n} e^{y_n})^{\alpha_n}$ for self-adjoint $x_i \in U_{\mathfrak{m}}$, skew-adjoint $y_i \in U_{\mathfrak{k}}$, and $\alpha_i = \pm 1$. Now $e^x e^y e^x \in \exp(\mathfrak{m})$ whenever $x, y \in \mathfrak{m}$ (Theorem 4.4), and inspection of the expression for $P(g) = gg^*$ shows that $P(g)$ lies in $\exp(\mathfrak{m})$ if $e^y e^x e^{-y} \in \exp(\mathfrak{m})$ whenever $x \in \mathfrak{m}$ and $y \in \mathfrak{k}$. Equivalently, we have to show that $\mathrm{Ad}(e^y)$ maps $\mathfrak{m}$ into $\mathfrak{m}$; since $\mathrm{Ad}(e^y) = e^{\mathrm{ad}(y)}$, it suffices to show that $\mathrm{ad}(y) : x \mapsto [y, x]$ maps $\mathfrak{m}$ into $\mathfrak{m}$, and this follows from the previous assertion. The set $M_G = \exp(\mathfrak{m})$ is a convex submanifold because $\mathfrak{m}$ is a closed Lie triple system (Corollary 4.5).

3. If $g \in G$, then $gg^* = e^{x_0}$ for some $x_0 \in \mathfrak{m}$. This implies that $|g| = e^{x_0/2} \in M_G \subset G$. Now we have $u_g = |g|^{-1} g \in G$, and clearly $u_g \in K$.

4. If $p \in M_G$, then $p = P(g_o) = g_o g_o^*$ for some $g_o \in G$. Then, if $g \in G$, $I_g(p) = gg_o g_o^* g^* = P(gg_o) \in M_G$. Note that I_g is an isometry of M_G, because M_G has the induced metric, so Lemma 2.5 applies.

5. If $x \in \mathfrak{m}$ and $u \in K$, then $e^x \in M_G$ hence $ue^x u^* = \exp(I_u(e^x)) \in M_G$. Hence $uxu^* = \ln(ue^x u^*) \in \mathfrak{m}$. Since $\langle I_u(y), x \rangle_2 = \langle y, I_{u^*}(x) \rangle_2$ (see Remark 2.1), we obtain the proof of the assertion concerning $\mathfrak{m}^{\perp}$.

Clearly I_p maps $T_q M_G$ isometrically onto $T_{I_p(q)} M_G$. Assume now $w \in T_q M_G{}^{\perp} = q^{1/2} \mathfrak{m}^{\perp} q^{1/2}$ (see Remark 4.13). If $u = (pqp)^{1/2} p^{-1} q^{-1/2}$, then $u \in G$ and $uu^* = 1$, hence $w_0 = u(q^{-1/2} w q^{-1/2}) u^* \in \mathfrak{m}^{\perp}$ by the previous assertion. Then $I_p(w) = pwp = (pqp)^{1/2} w_0 (pqp)^{1/2} \in T_{I_p(q)} M_G{}^{\perp}$.

6. The previous items show that $P : G \to M_G$ is surjective. Now $dP_g : g.\mathfrak{g} \to T_{gg^*} M_G$ is given by $g.x \mapsto I_g(x + x^*)$. Clearly this map has split kernel $g\mathfrak{k}$. Let $g = |g| u_g$ as above. For $z \in T_{gg^*} M_G$ we have, by Remark 4.13, $z = (gg*)^{1/2} w (gg*)^{1/2} = I_{|g|}(w)$ for some $w \in \mathfrak{m}$. Let $x = I_{u_g^*}(w/2)$, then $x \in \mathfrak{m} \subset \mathfrak{g}$ and $dP_g(gx) = z$. Hence the group $K = P^{-1}(1)$ is a submanifold of G because $P : G \to M_G$ is a submersion (Proposition 2.3 of Chapter II in [19]).

7. The map $T : M_G \times K \to G$ given by $T(p, u) = pu$ is clearly smooth and it is a bijection by the statements above. The inverse is given by $g \mapsto (|g|, |g|^{-1}g)$; since $|g| = \exp(\frac{1}{2} \ln(gg^*))$, the map T is a diffeomorphism. $\square$

Remark 4.19. For $M = \exp(\mathfrak{m})$ a convex closed manifold in Σ, consider $\mathfrak{g}_M = \mathfrak{m} \oplus [\mathfrak{m}, \mathfrak{m}]$. Then $\mathfrak{g}_M$ is a Banach–Lie subalgebra of $\mathcal{H}_{\mathbb{C}}$ due to the formal identity

$$[[x, y], [z, w]] + [z, [w, [x, y]]] + [w, [[x, y], z]] = 0$$

and the fact that $\mathfrak{m}$ is a Lie triple system. Let $G_M = \langle \exp(\mathfrak{g}_M) \rangle$. Then G_M is a connected Banach–Lie group with Banach–Lie algebra $\mathfrak{g}_M$ (Proposition 4.17). Since $(a + [b, c])^* = a + [c, b]$ for any $a, b, c \in \mathfrak{m}$, then $M \subset G_M$ and $G_M^* = G_M$. It is also clear that $\mathfrak{k} = [\mathfrak{m}, \mathfrak{m}]$ ($\mathfrak{k}$ as in Theorem 4.18). The elements of M are indeed the positive elements of G_M, and the elements of the stabilizer of 1 are the unitary operators of G_M. Note that G_M is a submanifold of $\mathcal{H}_{\mathbb{C}}^{\bullet}$ if and only if K is a submanifold of $\mathcal{U}(\mathcal{H}_{\mathbb{C}})$.

When $\mathfrak{m}$ is a commutative associative subalgebra, we have $\mathfrak{g}_M = \mathfrak{m}$ and also $G_M = M \subset \mathfrak{m}$ is an open set (in particular G_M is a submanifold of $\mathcal{H}_{\mathbb{C}}^\bullet$). In any case $\mathfrak{m} = Z(\mathfrak{m}) \oplus Z(\mathfrak{m})^\perp = \mathfrak{m}_0 \oplus \mathfrak{m}_1$ (here $Z(\mathfrak{m})$ denotes the set $\{x \in \mathfrak{m} : [x, y] = 0 \ \forall y \in \mathfrak{m}\}$), and $M = \exp(\mathfrak{m}_0)\exp(\mathfrak{m}_1) \simeq M_0 \times M_1$ where M_i are convex and closed, hence $G_M \simeq M_0 \times G_{M_1}$. Since $\langle x, [a, [b, c]]\rangle_2 = \langle c, [b, [a, x]]\rangle_2$ for any $a, b, c, x \in \mathfrak{m}$, it is easy to see that $Z(\mathfrak{m}) = [\mathfrak{m}, [\mathfrak{m}, \mathfrak{m}]]^\perp$.

The results above assert that, for a given convex submanifold $M = \exp(\mathfrak{m})$, we have $M_{G_M} = M$. On the other hand, for a given connected involutive Banach–Lie subgroup G, we have $G_{M_G} \subset G$, though in general $[\mathfrak{m}, \mathfrak{m}]$ can be strictly smaller than $\mathfrak{k}$, so the other inclusion does not necessarily hold. The equality holds iff $\mathfrak{k}$ is semi-simple, i.e. $[\mathfrak{k}, \mathfrak{k}] = \mathfrak{k}$ (equivalently, if $Z(\mathfrak{k}) = 0$).

It is a well-known result (see [16, p. 42]) that $\overline{[\mathsf{HS}, \mathsf{HS}]} = \mathsf{HS}$ and $\overline{[\mathsf{HS}^h, \mathsf{HS}^h]} = i\mathsf{HS}^h$. Therefore taking $\mathfrak{m} = \mathsf{HS}^h$, we get $\mathfrak{k} = i\mathsf{HS}^h$, and then $\mathfrak{g}_M = \mathsf{HS}$. This implies $G_{\Sigma/\mathbb{R}} = \mathcal{H}_{\mathbb{C}}^\bullet/\mathbb{C}^\times 1$. Clearly $P(G_\Sigma) = P(\mathcal{H}_{\mathbb{C}}^\bullet) = \Sigma$, because any positive invertible operator has an invertible square root. On the other hand it is clear that the isotropy group K equals $\mathcal{U}(\mathcal{H}_{\mathbb{C}})$ (the unitary group of $\mathcal{H}_{\mathbb{C}}$). So there is an analytic isomorphism given by polar decomposition: $\Sigma \simeq \mathcal{H}_{\mathbb{C}}^\bullet/\mathcal{U}(\mathcal{H}_{\mathbb{C}})$. The manifold of positive invertible operators Σ is an homogeneous manifold for the group of invertible operators $\mathcal{H}_{\mathbb{C}}^\bullet$, which acts isometrically and transitively on Σ. This last statement is well known, and Theorem 4.18 can be read as a natural generalization.

5. Projecting to closed convex submanifolds

We refer the reader to [19] for the first and second variation formulas.

Proposition 5.1. *Let M be a convex subset of Σ, and let $p \in \Sigma$. Then there is at most one normal geodesic γ of Σ joining p and M such that $L(\gamma) = \mathrm{dist}(p, M)$. In other words, there is at most one point $q \in M$ such that $\mathrm{dist}(p, q) = \mathrm{dist}(p, M)$.*

Proof. Suppose there are two such points, q and $r \in M$, joined by a geodesic $\gamma_3 \in M$, such that $L(\gamma_1) = \mathrm{dist}(p, q) = L(\gamma_2) = \mathrm{dist}(p, r) = d(p, M)$. We construct a proper variation of $\gamma \equiv \gamma_1$, which we call Γ_s. The construction follows the figure below, where $\sigma(s, t) := \sigma_s(t)$ is the geodesic joining p with $\gamma_3(s)$.
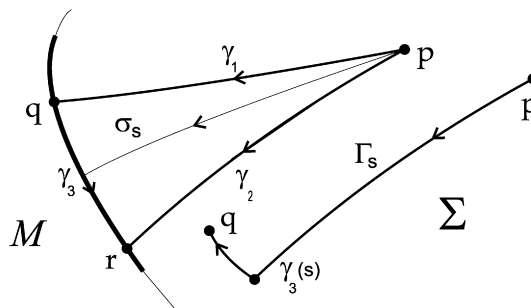
$$\Gamma(s, t) = \begin{cases} \sigma(s, t) & 0 \leqslant t \leqslant 1, \\ \gamma_3\,(s(2 - t)) & 1 \leqslant t \leqslant 2, \end{cases}$$

$$\gamma(t) = \Gamma(0, t) = \begin{cases} \gamma_1(t) & \text{if} \quad 0 \leqslant t \leqslant 1, \\ q & \text{if} \quad 1 \leqslant t \leqslant 2, \end{cases} \quad \text{so that} \quad \dot{\gamma}(t) = \begin{cases} \dot{\gamma}_1(t) & \text{if} \quad 0 \leqslant t \leqslant 1, \\ 0 & \text{if} \quad 1 \leqslant t \leqslant 2. \end{cases}$$

Also note that the variation vector field (which is a Jacobi field for γ) is given by

$$V(t) = \frac{\partial \Gamma}{\partial s}(t, 0) = \begin{cases} \frac{\partial \sigma}{\partial s}(t, 0) & 0 \leqslant t \leqslant 1, \\ (2 - t)\dot{\gamma}_3(0) & 1 \leqslant t \leqslant 2. \end{cases}$$

If $\Delta_i\dot\gamma$ denotes the jump of the tangent vector field to γ at t_i, namely $\dot\gamma(t_i^+) - \dot\gamma(t_i^-)$, and Γ is a proper variation of γ, then the first variation formula for the curve $\gamma : [0, 2] \to \Sigma$ reads

$$\|\dot\gamma\|_\gamma \frac{d}{ds}\bigg|_{s=0^+} L(\Gamma_s) = -\int_0^2 \langle V(t), D_t\dot\gamma(t)\rangle_{\gamma(t)}\, dt - \sum_{i=1}^{k-1}\langle V(t_i), \Delta_i\dot\gamma\rangle_{\gamma(t_i)}.$$

In this case, $D_t\dot\gamma$ is zero in the whole interval $[0, 2]$, because γ consists (piecewise) of geodesics. The jump points are $t_0 = 0$, $t_1 = 1$ and $t_2 = 2$, so the formula reduces to

$$\langle \dot\gamma_3(0), \dot\gamma_1(1)\rangle_q = \frac{d}{ds}\bigg|_{s=0^+} L(\Gamma_s)\|\dot\gamma\|_\gamma.$$

Recall that $\gamma_3 \subset M$, and that γ_1 is minimizing. Then the right-hand term is nonnegative, which proves that the angle between γ_1 and γ_3 at q is bigger that $\pi/2$. With a similar argument, we deduce that the same holds for the angle between γ_2 and γ_3 at r. Hence, the sum of the three inner angles of this geodesic triangle is *at least* π. Since the sum cannot exceed π (see Corollary 3.12), it follows that the angle subtended at p must be zero, which proves that γ_1 and γ_2 are the same geodesic, and uniqueness follows. □

Next we consider the problem of existence of the minimizing geodesic.

Proposition 5.2. *Let M be a convex submanifold of Σ, and p a point of Σ not in M. Then the existence of a geodesic β joining p with M such that $L(\beta) = \mathrm{dist}(p, M)$ is equivalent to the existence of a geodesic γ joining p with M with the property that γ is orthogonal to M.*

Proof. In fact, the existence of such a geodesic β is equivalent to the existence of a point $q \in M$ such that $\mathrm{dist}(p, M) = \mathrm{dist}(p, q)$. We will show that if $q \in M$ is a point such that γ_{qp} is orthogonal to M at q, then $\mathrm{dist}(q, p) = \mathrm{dist}(M, p)$. The other implication follows from the uniqueness theorem above. Consider the geodesic triangle generated by p, q and d, where d is any point in M different from q. Since γ_{qp} is orthogonal to T_qM, it is orthogonal to γ_{qd}. Then, by virtue of the Hyperbolic Cosine Law (Eq. (10) in Section 3), $\mathrm{dist}(d, p)^2 = L(\gamma_{dp})^2 \geqslant L(\gamma_{qp})^2 + L(\gamma_{qd})^2 > L(\gamma_{qp})^2 = \mathrm{dist}(q, p)^2$. □

This last proposition raises the following question: is the normal bundle NM of M diffeomorphic to Σ, *via* the exponential map?

Lemma 5.3. *Let M be a convex and closed submanifold. Let $E : NM \to \Sigma$ be the map $(q, v) \mapsto \mathrm{Exp}_q(v)$. For $\varepsilon > 0$, put $NM_\varepsilon = \{(p, v) \in NM : \|v\|_p < \varepsilon\}$ and $\Omega_\varepsilon = E(NM_\varepsilon)$. Then E is injective and there exists $\varepsilon > 0$ such that $E : NM_\varepsilon \to \Omega_\varepsilon$ is a C^ω-diffeomorphism. The set Ω_ε is an open neighborhood of M in Σ.*

Proof. Let us prove first that E is injective. Assume there exist $p, q \in M$, $v \in T_pM^\perp$, $w \in T_qM^\perp$ with $\mathrm{Exp}_p(v) = \mathrm{Exp}_q(w)$. Naming r to this point, consider the geodesic triangle in Σ spanned by $p, q \in M$ and $r \in \Sigma$. The geodesic which joins p to r is clearly $\gamma_1(t) = \mathrm{Exp}_p(tv)$, which is orthogonal to M at p, and the same is true for $\gamma_2(t) = \mathrm{Exp}_q(tw)$, which joins q to r. Hence $p = q$ and $v = w$ because of Corollary 3.12.

We may assume that $1 \in M$. Since $E(q, v) = qe^{q^{-1}v}$, the differential of E at $(1, 0) \in NM$ is the identity map because $T_1M \oplus T_1M^\perp = T_1\Sigma$ and $d\exp_0 = id$. The inverse mapping theorem ([19], Theorem 5.2 of Chapter I) gives C^ω-diffeomorphic neighborhoods $U_0 = \{(q, v) \in NM : \mathrm{dist}(q, 1) < \varepsilon, \|v\|_q < \varepsilon\} \subset NM_\varepsilon$ and $\Omega_0 = E(U_0) \subset \Sigma$ respectively. For given $(p, v) \in NM_\varepsilon$, consider the isometry of M given by $\tilde I_p : x \mapsto p^{1/2}xp^{1/2}$, and note that $\tilde I_p(1) = p$. If $(q, w) \in U_0$, then clearly $\tilde I_p(q) \in M$. Moreover, $\tilde I_p(w) \in T_{\tilde I_p(q)}M^\perp$ by Theorem 4.18, hence $U_p = (\tilde I_p \times \tilde I_p)(U_0)$ is an open neighborhood of (p, v) in NM_ε diffeomorphic to U_0. Now $E|_{U_p} : U_p \to E(U_p)$ is a diffeomorphism, because a straightforward computation shows that $E|_{U_p} = \tilde I_p \circ E \circ (\tilde I_p \times \tilde I_p)^{-1}$. □

Remark 5.4. Clearly $E(NM) \subset \Sigma$ is the set of points $p \in \Sigma$ with the following property: there is a point $q \in M$ such that $\mathrm{dist}(q, p) = \mathrm{dist}(M, p)$. Note that the map $\Pi_M : E(NM) \to M$, which assigns to $p \in E(NM)$ the unique point

$q \in M$ such that $\mathrm{dist}(q, p) = \mathrm{dist}(M, p)$, is surjective. This map is obtained *via* a geodesic that joins p and M, and this geodesic is orthogonal to M, therefore we call $\Pi_M(p)$ the *foot of the perpendicular* from M to p.

Lemma 5.5. *Let $p, q \in E(NM)$, and $\Pi_M(p) \neq \Pi_M(q)$. If γ_p is a geodesic that joins $\Pi_M(p)$ to p and γ_q is a geodesic that joins $\Pi_M(q)$ to q, put $f(t) = \mathrm{dist}(\gamma_p(t), \gamma_q(t))$. Then the map $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is increasing.*

Proof. Since f is a convex function (Corollary 3.10), it suffices to show that $f'(0^+) \geqslant 0$.

Take a variation $\sigma(t, s)$, where $\sigma_t(s)$ is the geodesic joining $\gamma_p(t)$ to $\gamma_q(t)$. Then $\sigma(t, 0) = \gamma_p(t)$, $\sigma(t, 1) = \gamma_q(t)$, and $\sigma(0, s) = \gamma(s)$ is the geodesic joining $\Pi_M(p)$ to $\Pi_M(q)$ (which is contained in M by virtue of the convexity). Note also that $\sigma(1, s)$ is the geodesic joining p to q. This construction is shown in the figure below.
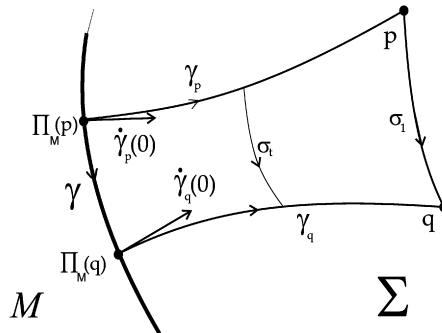
Note that $f(t) = L(\sigma_t)$. Put $V = \frac{d}{dt}\big|_{t=0}\, \sigma$. We apply the first variation formula to obtain

$$\|\dot{\gamma}\|_\gamma \frac{d}{dt}\bigg|_{t=0^+} L(\sigma_t) = -\int_0^1 \langle V(s), D_s\dot{\gamma}(s)\rangle_{\gamma(s)}\, ds + \langle V(1), \dot{\gamma}(1)\rangle_{\Pi_M(p)} - \langle V(0), \dot{\gamma}(0)\rangle_{\Pi_M(q)}.$$

The fact that γ is a geodesic reduces the formula to

$$\|\dot{\gamma}\|_\gamma f'(0^+) = -\langle V(1), -\dot{\gamma}(1)\rangle_{\Pi_M(p)} + \langle -V(0), \dot{\gamma}(0)\rangle_{\Pi_M(q)}.$$

Note also that $V(0) = \dot{\gamma}_p(0)$, $V(1) = \dot{\gamma}_q(0)$. Recalling that the angles at M are right angles, we obtain $f'(0^+) = 0$. $\quad\square$

Theorem 5.6. *The map Π_M is a contraction, namely $\mathrm{dist}(\Pi_M(p), \Pi_M(q)) \leqslant \mathrm{dist}(p, q)$.*

Proof. We may assume again that $p, q \notin M$, and that $\Pi_M(p) \neq \Pi_M(q)$. In the notation of the lemma above, note that $f(0) = d(\Pi_M(p), \Pi_M(q))$ and $f(1) = \mathrm{dist}(p, q)$; since f is increasing, the assertion is proved. $\quad\square$

We want to prove that $E(NM) = \Sigma$. We will do this by proving that it is both open and closed in Σ. The following argument is similar to the one used by H. Porta and L. Recht in [25].

Lemma 5.7. *For $\lambda \in [1, +\infty)$, put $\eta_\lambda : E(NM) \to E(NM)$, $\eta_\lambda(\mathrm{Exp}_p(v)) = \mathrm{Exp}_p(\lambda v)$. Let Ω_ε be as in Lemma 5.3. Then $E(NM) = \bigcup_{\lambda \geqslant 1} \eta_\lambda(\Omega_\epsilon)$, and each $\eta_\lambda : \Omega_\epsilon \to \Sigma$ is a C^ω diffeomorphism onto its open image.*

Proof. Clearly $\bigcup_{\lambda \geqslant 1} \eta_\lambda(\Omega_\epsilon) \subset E(NM)$. Let us prove the other inclusion. First, if $r = \mathrm{Exp}_p(v)$ with $\|v\|_p < \epsilon$ then $r \in \Omega_\epsilon = \eta_1(\Omega_\epsilon)$. Let us consider the case where $\|v\|_p \geqslant \epsilon$; then $r = \mathrm{Exp}_p(v) = \mathrm{Exp}_p(\lambda w)$ with $\lambda = \frac{2\|v\|_p}{\varepsilon}$ and $w = \frac{\varepsilon}{2\|v\|_p} v$, so $r \in \eta_\lambda(\Omega_\varepsilon)$ because $\|w\|_p = \varepsilon/2 < \varepsilon$ and $\lambda \geqslant 1$.

Assume that there exist $r_1, r_2 \in \Omega_\varepsilon$ and $\lambda \geqslant 1$ such that $\eta_\lambda(r_1) = \eta_\lambda(r_2)$. That is, assume there exist $p, q \in M$, $v \in T_pM^\perp$, $w \in T_qM^\perp$ with $\|v\|_p < \varepsilon$, $\|w\|_q < \varepsilon$ and $\mathrm{Exp}_p(\lambda v) = \mathrm{Exp}_q(\lambda w)$, namely $E(p, \lambda v) = E(q, \lambda w)$. Since E is injective by Lemma 5.3, we have $p = q$ and $v = w$. This argument proves that the maps η_λ are injective.

Next we show that, for any $\lambda \geqslant 1$ and $r \in \Omega_\epsilon$, $d(\eta_\lambda)_r : T_r \Sigma \to T_{\eta_\lambda(r)} \Sigma$ is a linear isomorphism, and this will prove the final assertion. Take $\alpha \subset \Omega_\epsilon$ a geodesic such that $\alpha(0) = r$ and $\dot{\alpha}(0) = x$. Since α is a geodesic, we have that $\text{dist}(\alpha(t), r) = t \|\dot{\alpha}(0)\|_r$ for $t \geqslant 0$ (see Section 3.4). Put $\beta = \eta_\lambda \circ \alpha$. Then $\beta(0) = \eta_\lambda(r)$ and $\dot{\beta}(0) = d(\eta_\lambda)_r(x)$. Clearly $\text{dist}(\beta(t), \eta_\lambda(r)) \leqslant L_0^t(\beta) = \int_0^t \|\dot{\beta}(s)\|_{\beta(s)} \, ds$. On the other hand, $\text{dist}(\eta_\lambda(\alpha(t)), \eta_\lambda(r)) \geqslant \text{dist}(\alpha(t), r) = t \|x\|_r$ where the inequality is due to Lemma 5.5, because $\lambda \geqslant 1$. If we put together these two inequalities and divide by t, we get

$$\frac{1}{t} \int_0^t \|\dot{\beta}(s)\|_{\beta(s)} \, ds \geqslant \|x\|_r.$$

Taking limit for $t \to 0^+$ gives $\|d(\eta_\lambda)_r(x)\|_{\eta_\lambda(r)} \geqslant \|x\|_r$. Now put $A_\lambda = \tilde{I}_{\eta_\lambda(r)}^{-1} \circ d(\eta_\lambda)_r \circ \tilde{I}_r$, where $\tilde{I}_p : v \mapsto p^{1/2} v p^{1/2}$ are linear isomorphisms (see Lemma 2.5). If we consider $A_\lambda : T_1 \Sigma \to T_1 \Sigma = \mathcal{H}_{\mathbb{R}}$, the last inequality says that $\|A_\lambda(x)\|_2 \geqslant \|x\|_2$ for any $x \in \mathcal{H}_{\mathbb{R}}$.

Clearly $\eta_1 = id_{\Omega_\epsilon}$ and $d(\eta_1)_r = id_{T_r \Sigma}$. Since the map $(\lambda, r) \mapsto \eta_\lambda(r)$ is analytic from $\mathbb{R}_{)0} \times \Omega_\epsilon$ to Σ, there is an open neighborhood of $1 \in \mathbb{R}$ such that A_λ is an isomorphism. Assume A_λ is invertible for $\lambda \in [1, m)$: then $\|A_\lambda^{-1}\|_{B(\mathcal{H}_{\mathbb{R}})} \leqslant 1$ for any $\lambda \in [1, m)$. Since $A_m = \lim_{\lambda \to m^-} A_\lambda$ (in the operator norm of $B(\mathcal{H}_{\mathbb{R}})$) and $\|A_m A_\lambda^{-1} - 1\| \leqslant \|A_m - A_\lambda\| < 1$ if λ is close enough to m, it follows that $A_m A_\lambda^{-1}$ is invertible, thus A_m is invertible. Since the maps $\tilde{I}_p$ are isomorphisms, $d(\eta_\lambda)_r$ is an isomorphism for any $\lambda \geqslant 1$, and any $r \in \Omega_\epsilon$.

Corollary 5.8. *The set $E(NM)$ is open in Σ.*

Theorem 5.9. *Let M be a convex closed submanifold of Σ. Then for every point $p \in \Sigma$, there is a unique normal geodesic γ_p joining p to M such that $L(\gamma_p) = \text{dist}(p, M)$. This geodesic is orthogonal to M, and if $\Pi_M : \Sigma \to M$ is the map that assigns to p the end-point of γ_p, then Π_M is a contraction for the geodesic distance.*

Proof. The theorem will follow if we prove that $E(NM) = \Sigma$. Since Σ is connected and $E(NM)$ is open, it suffices to prove that $E(NM)$ is also closed. Let $p \in \overline{E(NM)}$. There exist points $q_n \in M$, $v_n \in T_{q_n} M^\perp$ such that $p = \lim_n p_n = \lim_n \text{Exp}_{q_n}(v_n)$. Observe that $q_n = \Pi_M(p_n)$, so $\text{dist}(q_n, q_m) \leqslant \text{dist}(p_n, p_m)$. Since $\{p_n\}$ converges to p, it is a Cauchy sequence. It follows that $\{q_n\}$ is also a Cauchy sequence. Since M is closed (and therefore complete), there exists $q \in M$ such that $q = \lim_n q_n$. We assert that $\text{dist}(p, q) = \text{dist}(p, M)$. First note that $\text{dist}(p, q_n) \leqslant \text{dist}(p, p_n) + \text{dist}(p_n, q_n)$ and $\text{dist}(p_n, q_n) = \text{dist}(p_n, M)$, so $\text{dist}(p, q_n) \leqslant \text{dist}(p, p_n) + \text{dist}(p_n, M)$. Taking limits gives $\text{dist}(p, q) \leqslant \text{dist}(p, M)$. $\quad \square$

Note that Σ decomposes as a direct product: with the contraction Π_M, we can decompose Σ by picking, for fixed p,

1. the unique point $q = \Pi_M(p)$ such that $\text{dist}(p, q) = \text{dist}(p, M)$;
2. a vector v_p normal to $T_q M$ such that the geodesic in Σ with initial velocity v_p starting at q passes through p; note that $v_p = \text{Exp}_{\Pi_M(p)}^{-1}(p)$, and also $\|v_p\|_q = \text{dist}(p, M)$.

Since the exponential map is analytic on both of its variables, we get

Theorem 5.10. *The map $p \mapsto (\Pi_M(p), v_p)$ is the inverse of the map $(q, v_q) \mapsto \text{Exp}_q(v_q)$, and gives a real-analytic isomorphism between the manifolds Σ and NM.*

Theorem 5.11. *Fix a closed convex submanifold M of Σ. Let $a \in \Sigma$. Then there exist unique operators $c \in \Sigma$, $v \in \mathcal{H}_{\mathbb{R}}$ such that $c \in M$, $v \in T_c M^\perp$, and $a = c e^{c^{-1} v}$.*

Using the tools of Section 4, we can write the factorization theorem in terms of intrinsic operator equations (see [22] for the finite dimensional analogue):

Theorem 5.12. *Assume* $\mathfrak{m} \subset \mathcal{H}_\mathbb{R}$ *is a Lie triple system. Then for any operator* $a \in \mathcal{H}_\mathbb{R}$*, there exist unique operators* $x \in \mathfrak{m}$ *and* $v \in \mathfrak{m}^\perp$ *such that the following decomposition holds:* $\mathrm{e}^a = \mathrm{e}^x \mathrm{e}^v \mathrm{e}^x$*. The map* $d_a : \mathcal{H}_\mathbb{R} \to \mathbb{R}$*,* $d_a(y) = \|\ln(\mathrm{e}^{a/2} \mathrm{e}^{-y} \mathrm{e}^{a/2})\|_2$ *has the operator* $2x$ *as its unique minimizer in* $\mathfrak{m}$*.*

As a corollary, we obtain a polar decomposition relative to a convex submanifold.

Theorem 5.13. *Assume* $M = \exp(\mathfrak{m}) \subset \Sigma$ *is a closed convex submanifold. Then for any* $g \in \mathcal{H}_\mathbb{C}^\bullet$ *there is a unique factorization of the form* $g = \mathrm{e}^x \mathrm{e}^v u$ *where* $x \in \mathfrak{m}$*,* $v \in \mathfrak{m}^\perp$ *and* $u \in \mathcal{U}(\mathcal{H}_\mathbb{C})$ *is a unitary operator. The map* $g \mapsto (\mathrm{e}^x, \mathrm{e}^v, u)$ *is an analytic bijection which gives an isomorphism*

$$\mathcal{H}_\mathbb{C}^\bullet \simeq M \times \exp(\mathfrak{m}^\perp) \times \mathcal{U}(\mathcal{H}_\mathbb{C}).$$

Proof. Since $gg^* \in \Sigma$, we can write $gg^* = \mathrm{e}^x \mathrm{e}^{2v} \mathrm{e}^x$ with $x \in \mathfrak{m}$ and $v \in \mathfrak{m}^\perp$. If $u = (\mathrm{e}^x \mathrm{e}^v)^{-1} g = \mathrm{e}^{-v} \mathrm{e}^{-x} g$ we have $uu^* = \mathrm{e}^{-v} \mathrm{e}^{-x} gg^* \mathrm{e}^{-x} \mathrm{e}^{-v} = 1$ and also $u^* u = g^* \mathrm{e}^{-x} \mathrm{e}^{-v} \mathrm{e}^{-v} \mathrm{e}^{-x} g = 1$. Hence u is a unitary operator and $g = \mathrm{e}^x \mathrm{e}^v u$. This factorization is unique because if $g = \mathrm{e}^{x_1} \mathrm{e}^{v_1} u_1 = \mathrm{e}^{x_2} \mathrm{e}^{v_2} u_2$, then $gg^* = \mathrm{e}^{x_1} \mathrm{e}^{2v_1} \mathrm{e}^{x_1} = \mathrm{e}^{x_2} \mathrm{e}^{2v_2} \mathrm{e}^{x_2}$, so $x_1 = x_2$, $v_1 = v_2$ and then $u_1 = u_2$. $\square$

6. Projecting to the manifold of diagonal operators

Lemma 6.1. *Let* $\alpha, \beta \in \mathbb{R}$ *and* $a, b \in \mathsf{HS}^h$*. Then*

$$\mathrm{Exp}_{\alpha+a}(\beta + b) = \alpha \mathrm{e}^{\beta/\alpha} + k$$

where k *is a self-adjoint Hilbert–Schmidt operator.*

Proof. It is a straightforward computation:

$$
\begin{aligned}
(\alpha + a)\mathrm{e}^{(\alpha+a)^{-1}(\beta+b)} &= (\alpha + a)\big[1 + (\alpha + a)^{-1}(\beta + b) + \cdots\big] \\
&= (\alpha + a)\Big[1 + \beta/\alpha + \tfrac{1}{2}(\beta/\alpha)^2 + \cdots + k\Big]. \qquad \square
\end{aligned}
$$

We need some remarks before we proceed. Fix an orthonormal basis $\{e_i\}_{i \in \mathbb{N}}$ of H.

1. Consider the diagonal manifold $\Delta \subset \Sigma$:

   $$\Delta = \{d + \alpha > 0: \ \alpha \in \mathbb{R}, \ d \text{ is a diagonal Hilbert–Schmidt operator}\}.$$

   It is closed and geodesically convex. This is due to the fact that the diagonal operators form a closed commutative associative subalgebra.
2. If $d_0 \in \Delta$, then $T_{d_0}\Delta = \{\alpha + d; \ \alpha \in \mathbb{R}, \ d \in \mathsf{HS} \text{ is diagonal and real}\} = T_1\Delta$ (see Remark 4.13).
3. Consider the map $A \mapsto A^D =$ the diagonal part of A. Then
   (a) For Hilbert–Schmidt operators we have $A^D = \sum_i p_i A p_i$ where convergence is in the 2-norm (and hence in the operator norm); here $p_i = e_i \otimes e_i = \langle e_i, \cdot \rangle e_i$ is the orthogonal projection onto the real line generated by e_i.
   (b) $(A^D)^D = A^D$ and $\mathrm{tr}(A^D A) = \mathrm{tr}((A^D)^2)$.
   (c) $\mathrm{tr}(A^D B) = \mathrm{tr}(AB)$ if B is diagonal.
4. The scalar manifold $\Lambda = \{\lambda 1: \ \lambda \in \mathbb{R}_{>0}\}$ is convex and closed in Σ, with tangent space at any $\lambda \in \Lambda$ given by $\mathbb{R}1 \subset \mathcal{H}_\mathbb{R}$.
5. A vector $v = \mu + u$ is contained in $T_{d_0}\Delta^\perp$ if and only if $\mu = 0$ and $u^D = 0$. This follows from Remark 4.13, the fact that $\mu + u^D \in T_{d_0}\Delta$, and Remark (3) of this list. In other words for any $d_0 \in \Delta$,

   $$T_{d_0}\Delta^\perp = T_1\Delta^\perp = \{v \in \mathsf{HS}^h: \ v \text{ is codiagonal}\} =: \Gamma.$$

Theorem 6.2. *Let* $a \in \mathsf{HS}^h$*. Then there exist* $\lambda \in \mathbb{R}_{>0}$*,* $d \in \Delta$ *and* $x \in \mathsf{HS}^h$ *such that*:

$$a + \lambda = (d + \lambda)\mathrm{e}^{(d+\lambda)^{-1}v} = (d + \lambda)^{\frac{1}{2}} \mathrm{e}^{(d+\lambda)^{-\frac{1}{2}} v (d+\lambda)^{-\frac{1}{2}}} (d + \lambda)^{\frac{1}{2}}.$$

Moreover, for fixed λ, d and v are unique and a + λ ↦ (d, v) (which maps Σ → NΔ) is a real analytic isomorphism between manifolds.

Proof. Let $\lambda = \|a\|_\infty + \epsilon$, for any $\epsilon > 0$. Then $p = a + \lambda \in \Sigma$. Let $\Pi_\Delta(p) = d + \alpha$, where $D \in \Delta$. Now pick the unique $v \in T_{d+\alpha}\Delta^\perp$ such that $\mathrm{Exp}_{d+\alpha}(v) = p$, this operator v has the desired form because of remark (5) above. As a consequence of Lemma 6.1 $\alpha = \lambda$, for in this case $\beta = 0$. □

This theorem can be rephrased saying that, given a self-adjoint Hilbert–Schmidt operator a, for any $\lambda \in \mathbb{R}_{>0}$ such that $a + \lambda > 0$, one has a unique factorization $a + \lambda = De^w D$ where $D = (\lambda + d)^{1/2} > 0$ is a diagonal operator and $w = D^{-1}vD^{-1} \in \Gamma$ is a self-adjoint operator with null diagonal. The normal bundle clearly splits in this case, so

Proposition 6.3. *Consider the submanifolds Δ, $\exp(\Gamma) \subset \Sigma$. Then the projection map Π_Δ induces a diffeomorphism $\Sigma \simeq \Delta \times \exp(\Gamma)$.*

Corollary 6.4. *For any $g \in \mathcal{H}_\mathbb{C}^\bullet$, there is a unique factorization $g = de^w u$, where d is a positive invertible diagonal operator of $\mathcal{H}_\mathbb{C}$, w is a self-adjoint operator with null diagonal in $\mathcal{H}_\mathbb{C}$ and u is a unitary operator of $\mathcal{H}_\mathbb{C}$.*

Proof. The previous results together with Theorem 5.13. □

7. A foliation of codimension one

In this section we describe a foliation of the total manifold, and show how to translate the results from previous sections to a particular leaf (the submanifold Σ_1) in order to show an application concerning (finite dimensional) matrix algebras. Recall that we write HS^h for the self-adjoint Hilbert–Schmidt operators. Fix $\lambda \in \mathbb{R}_{>0}$. Let

$$\Sigma_\lambda = \{a + \lambda \in \Sigma, \ a \in \mathsf{HS}^h\}.$$

Observe that $\Sigma_\lambda \cap \Sigma_\beta = \emptyset$ when $\lambda \neq \beta$, since $a + \lambda = b + \beta$ implies $a - b = \beta - \lambda$. In this way, we can decompose the total space by means of these leaves, $\Sigma = \bigcup_{\lambda > 0}^{\cdot} \Sigma_\lambda$.

Proposition 7.1. *The leaves Σ_λ are geodesically convex closed submanifolds.*

Proof. We consider the projection to the convex scalar manifold Λ (see remark (4) above). The fact that the projection Π_Λ is a contraction (therefore a continuous map) implies that Σ_λ is closed; one must only observe that $\Sigma_\lambda = \Pi_\Lambda^{-1}(\lambda)$. To show that Σ_λ is geodesically convex we recall that, by virtue of Lemma 6.1, for any real $\lambda > 0$ and any $p \in \Sigma_\lambda$, there is an identification *via* the inverse exponential map at p, $T_p\Sigma_\lambda = \mathsf{HS}^h$. □

Remark 7.2. Take $\delta + c \in T_{a+\lambda}\Sigma_\lambda^\perp$. Since $T_{a+\lambda}\Sigma_\lambda$ can be identified with HS^h, the equality $\langle \delta + c, d \rangle_{a+\lambda} = 0 \ \forall d \in \mathsf{HS}^h$ is equivalent to

$$\mathrm{tr}\big[(a + \lambda)^{-1}\big[(\delta + c)(a + \lambda)^{-1} - \delta/\lambda\big]d\big] = 0 \quad \forall d \in \mathsf{HS}^h.$$

Equivalently, $T_{a+\lambda}\Sigma_\lambda^\perp = \mathrm{span}(a + \lambda)$; shortly $T_p\Sigma_\lambda^\perp = \mathrm{span}(p)$ for any $p \in \Sigma_\lambda$.

Proposition 7.3. *Fix real $\alpha, \lambda > 0$. Let $\Pi_{\alpha,\lambda} = \Pi_{\Sigma_\lambda}|_{\Sigma_\alpha} : \Sigma_\alpha \to \Sigma_\lambda$. Then*

1. *$\Pi_{\alpha,\lambda}(p) = \frac{\lambda}{\alpha}p$, so $\Pi_{\alpha,\lambda}(p)$ commutes with p.*
2. *$\Pi_{\alpha,\lambda}$ is an isometric bijection between Σ_α and Σ_λ, with inverse $\Pi_{\lambda,\alpha}$.*
3. *$\Pi_{\alpha,\lambda}$ gives parallel translation along vertical geodesics joining both leaves (that is, geodesics orthogonal to both leaves).*

Proof. Notice that for a point $b + \alpha \in \Sigma_\alpha$ to be the endpoint of the geodesic γ, starting at $a + \lambda \in \Sigma_\lambda$, such that $L(\gamma) = \mathrm{dist}(b + \alpha, \Sigma_\lambda)$, we must have

$$b + \alpha = \mathrm{Exp}_{a+\lambda}(x + c) = \mathrm{Exp}_{a+\lambda}\big(k.(a + \lambda)\big) = e^k(a + \lambda)$$
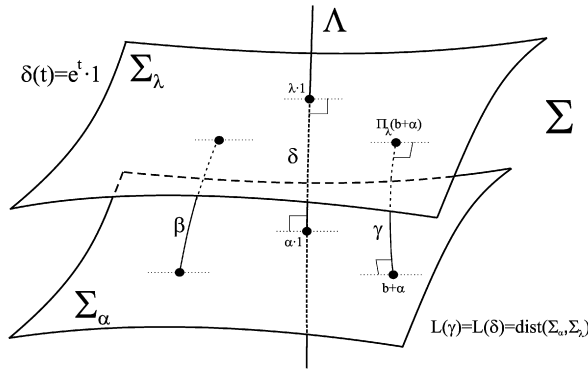
Fig. 1. The geodesics γ and δ are minimizing, the geodesic β is not.

where $k \in \mathbb{R}$ comes from Remark 7.2 above, since $x + c \in T_{a+\lambda}\Sigma_\lambda^\perp$. From Lemma 6.1, we deduce that $k = \ln(\frac{\alpha}{\lambda})$, and $a = \frac{\lambda}{\alpha}b$. So, $b + \alpha = \frac{\alpha}{\lambda}(a + \lambda)$ and also $\gamma(t) = (a + \lambda)(\frac{\alpha}{\lambda})^t$. Now it is clear that $\Pi_\lambda(b + \alpha) = \frac{\lambda}{\alpha}(b + \alpha)$ commutes with $b + \alpha$. To prove that Π is isometric, observe that

$$\text{dist}\big(\Pi_{\alpha,\lambda}(p), \Pi_{\alpha,\lambda}(q)\big) = \left\| \ln\left(\left(\frac{\lambda}{\alpha}p\right)^{-\frac{1}{2}} \left(\frac{\lambda}{\alpha}q\right) \left(\frac{\lambda}{\alpha}p\right)^{-\frac{1}{2}} \right) \right\|_2 = \left\| \ln(p^{-\frac{1}{2}}qp^{-\frac{1}{2}}) \right\|_2 = \text{dist}(p, q).$$

That Π gives parallel translation along γ follows from the formula for Π given in the first item of this proposition and the formula for the parallel translation given in Remark 4.14.

The normal bundle in the case of $M = \Sigma_1$ can be thought of as a direct product:

Proposition 7.4. *The map* $T : \Sigma \to \Sigma_1 \times \Lambda$, *which assigns* $a + \alpha \mapsto (\frac{1}{\alpha}(a + \alpha), \alpha)$ *is bijective and isometric* (Σ_1 *and* Λ *have the induced submanifold metric*). *In other words, there is a Riemannian isomorphism* $\Sigma \simeq \Sigma_1 \times \Lambda$.

Proposition 7.5. *The leaves* Σ_α, Σ_λ *are also parallel in the following sense*: *any minimizing geodesic joining a point in one of them with its projection in the other is orthogonal to both of them. See Fig.* 1.

For any $b + \alpha \in \Sigma_\alpha$ *we have* $\text{dist}(b + \alpha, \Sigma_\lambda) = \text{dist}(\Sigma_\alpha, \Sigma_\lambda) = |\ln(\frac{\alpha}{\lambda})|$. *In particular, the distance between* α, λ *in the scalar manifold* Λ *is given by the Haar measure of the open interval* (α, β) *on the multiplicative group* $\mathbb{R}_{>0}$.

Proof. It is a straightforward computation that follows from the previous results; the last statement was observed by E. Vesentini in another context [27]. □

Since Σ is a symmetric space, curvature is preserved when we parallel-translate bidimensional planes. Note also that vertical planes (i.e., planes generated by a vector $v \in \mathsf{HS}^h = T_\lambda\Sigma_\lambda$ and λ) are commuting sets of operators.

Proposition 7.6. *Let* $p \in \Sigma_\lambda$. *Then the sectional curvature of vertical* 2-*planes is zero.*

Proof. It follows from the formula for the curvature given in Section 3.1. □

7.1. The embedding of M_n^+ *in* Σ_1

Let M_n^+ be the set of positive invertible $n \times n$ matrices (see the introduction of this paper). First note that we can embed $M_n^+ \hookrightarrow \Sigma_1$ for any $n \in \mathbb{N}$: fix an orthonormal basis $\{e_n\}_{n \in \mathbb{N}}$ of H, let $p_{ij} = e_i \otimes e_j$, and identify the set M_n of real $n \times n$ matrices with the set

$$\mathcal{T} = \left\{ \sum_{i,j=1}^n a_{ij}\, p_{ij} \colon a_{ij} = a_{ji} \in \mathbb{R} \right\} \subset \mathsf{HS}^h.$$

We identify the manifold M_n^+ with $\mathcal{P} = \{e^T : T \in \mathcal{T}\} \subset \Sigma_1$ and the tangent space at each $e^T \in \mathcal{P}$ is $\mathcal{T}$. The set $\mathcal{P}$ is closed and convex in Σ_1 by Corollary 4.5. Let us call $S = \text{span}(e_1, \ldots, e_n)$, $S^\perp = \text{span}(e_{n+1}, e_{n+2}, \ldots)$. The operator P_S is the orthogonal projection to S and $Q_S = 1 - P_S$ is the orthogonal projection to $S^\perp$. Using matrix blocks, for any operator $A \in \text{B}(S)$, we identify

$$\mathcal{T} = \left\{ \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \right\} \quad \text{and} \quad \mathcal{P} = \left\{ \begin{pmatrix} e^A & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Remark 7.7. There is a direct sum decomposition of $\text{HS}^h = \mathcal{T} \oplus \mathcal{J}$ where operators in $J \in \mathcal{J}$ are such that $P_S J P_S = 0$. A straightforward computation using the matrix-block representation shows that $\langle a, b \rangle_2 = 0$ for any $a \in \mathcal{T}, b \in \mathcal{J}$, which says $\mathcal{T}^\perp = \mathcal{J}$ (here we consider HS^h as the total space). So the manifolds $\exp(\mathcal{J})$ and $\mathcal{P} = \exp(\mathcal{T})$ are orthogonal at 1, the unique intersection point. In the notation of Theorem 4.18, it is also clear that $\mathcal{P} = M_G \simeq G/K$, where

$$G = \begin{pmatrix} GL(n, \mathbb{C}) & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} \mathcal{U}(n, \mathbb{C}) & 0 \\ 0 & 1 \end{pmatrix}.$$

Theorem 7.8. *Let $\mathcal{P} \simeq M_n^+ \subset \Sigma_1$ with the above identification. Then for any positive invertible operator $e^b \in \Sigma_1$, ($b \in \text{HS}^h$) there is a unique factorization of the form*

$$e^b = \begin{pmatrix} e^A & 0 \\ 0 & 1 \end{pmatrix} \exp \left\{ \begin{pmatrix} e^{-A} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & Y^* \\ Y & X \end{pmatrix} \right\} \quad \text{where if } a = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{T}$$

then $e^a = e^A P_S + Q_S \in \mathcal{P} \simeq M_n^+$, $X^* = X$ is a Hilbert–Schmidt operator acting on the Hilbert space $S^\perp$, and $Y \in \text{B}(S, S^\perp)$.

An equivalent expression for the factorization is

$$e^b = \begin{pmatrix} e^{A/2} & 0 \\ 0 & 1 \end{pmatrix} \exp \left\{ \begin{pmatrix} 0 & e^{-A/2} Y^* \\ Y e^{-A/2} & X \end{pmatrix} \right\} \begin{pmatrix} e^{A/2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Proof. From previous theorems and the observations we made, we know that $e^b = e^{a/2} C e^{a/2}$, where

$$C = \exp \left\{ \begin{pmatrix} e^{-A/2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_{11} & V_{21}^* \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} e^{-A/2} & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

for some $A \in \text{B}(S)$ and some $v \in \text{HS}^h$. That $V_{11} = 0$ follows from the fact (see Remark 7.7) that $\mathcal{T}^\perp = \mathcal{J}$, and $v \in T_{e^a} \mathcal{P}^\perp$ if and only if $\text{tr}(e^{-A} B e^{-A} V_{11}) = 0$ for any $B \in \mathcal{T}$. $\square$

Remark 7.9. For any $b \in \text{HS}^h$, the operator

$$e^a = e^A P_S + P_{S^\perp} = \begin{pmatrix} e^A & 0 \\ 0 & 1 \end{pmatrix} = \exp \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$$

is the 'first block' $n \times n$ matrix which is closest to e^b in Σ, and with a slight abuse of notation for the traces of $\text{B}(S)$ and $\text{B}(S^\perp)$, we have

$$\text{dist}(\mathcal{P}, e^b) = \text{dist}(e^a, e^b) = \sqrt{\|Y e^{-A/2}\|_2^2 + \|X\|_2^2}.$$

Corollary 7.10. *For any $g \in \mathcal{H}_\mathbb{C}^\bullet$ there is a unique factorization $g = \lambda r e^v u$, where $\lambda \in \mathbb{R}_{>0}$, $u \in \mathcal{U}(\mathcal{H}_\mathbb{C})$ is a unitary operator,*

$$r = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}, \qquad v = \begin{pmatrix} 0 & Y^* \\ Y & X \end{pmatrix}$$

with $R \in \text{B}(S)^+ \simeq M_n^+$, $X = X^* \in \text{B}(S^\perp)$ a Hilbert–Schmidt, and $Y \in \text{B}(S, S^\perp)$.

Proof. We use the notation of Remark 7.7. Note that, by Theorem 5.13, $g = re^x u$ with $u \in \mathcal{U}(\mathcal{H}_\mathbb{C})$, $r \in \mathcal{P} = \exp(\mathcal{T})$ and $x \in \mathcal{T}^\perp$. But $\mathcal{T}^\perp = \mathcal{J} \oplus \mathbb{R}1$ if we consider $\mathcal{H}_\mathbb{R}$ as the total space, and $e^{\alpha+a} = e^\alpha e^a$ if $\alpha \in \mathbb{R}$. □

Acknowledgement

References

[1] E. Andruchow, G. Corach, M. Milman, D. Stojanoff, Geodesics and interpolation, Revista de la Unión Matemática Argentina 40 (3–4) (1997) 83–91.

[2] E. Andruchow, G. Larotonda, Nonpositively curved metric in the positive cone of a finite von Neumann algebra, J. London Math. Soc., in press.

[3] C.J. Atkin, The Hopf–Rinow theorem is false in infinite dimensions, Bull. London Math. Soc. 7 (1975) 261–266.

[4] C.J. Atkin, Geodesic and metric completeness in infinite dimensions, Hokkaido Math. J. 26 (1997) 1–61.

[5] W. Ballmann, Spaces of nonpositive curvature, Jahresber. Deutsch. Math. Verein. 103 (2) (2001) 52–65.

[6] R. Bhatia, On the exponential metric increasing property, Linear Algebra Appl. 375 (2003) 211–220.

[7] P. Calderón, Intermediate spaces and interpolation, the complex method, Studia Math. 24 (1964) 113–190.

[8] G. Corach, A. Maestripieri, Differential and metrical structure of positive operators, Positivity 3 (1999) 297–315.

[9] G. Corach, A. Maestripieri, Positive operators on Hilbert space: A geometrical view point, in: Colloquium on Homology and Representation Theory, Bol. Acad. Nac. Cienc. (Córdoba) 65 (2000) 81–94.

[10] G. Corach, H. Porta, L. Recht, Differential geometry of spaces of relatively regular operators, Integral Equations Operator Theory 13 (1990) 771–794.

[11] G. Corach, H. Porta, L. Recht, Splitting of the positive set of a C*-algebra, Indag. Math. NS 2 (4) (1991) 461–468.

[12] G. Corach, H. Porta, L. Recht, The geometry of the space of selfadjoint invertible elements in a C*-algebra, Integral Equations Operator Theory 16 (1993) 333–359.

[13] P. Eberlein, Geometry of Nonpositively Curved Manifolds, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1996.

[14] N. Grossman, Hilbert manifolds without epiconjugate points, Proc. AMS 16 (1965) 1365–1371.

[15] S. Helgason, Differential Geometry, Lie Groups and Symmetric Spaces, Academic Press, New York, 1962.

[16] P. de la Harpe, Classical Banach–Lie Algebras and Banach–Lie Groups of Operators in Hilbert Space, Lecture Notes in Mathematics, vol. 285, Springer, Berlin, 1972.

[17] H. Jost, Nonpositive Curvature: Geometric and Analytic Aspects, Lectures in Mathematics, Birkhäuser, Berlin, 1997.

[18] S. Lang, Introduction to Differentiable Manifolds, Interscience, New York, 1962.

[19] S. Lang, Differential and Riemannian Manifolds, Springer-Verlag, Berlin, New York, 1995.

[20] J. Lawson, Y. Lim, Symmetric spaces with convex metrics, preprint, 2006.

[21] J. McAlpin, Infinite dimensional manifolds and Morse theory, PhD thesis, Columbia University, 1965.

[22] G.D. Mostow, Some new decomposition theorems for semi-simple groups, Mem. Amer. Math. Soc. 14 (1955) 31–54.

[23] K.H. Neeb, A Cartan–Hadamard theorem for Banach-Finsler manifolds, Geom. Ded. 95 (2002) 115–150.

[24] H. Porta, L. Recht, Spaces of projections in Banach algebras, Acta Matemática Venezolana 38 (1987) 408–426.

[25] H. Porta, L. Recht, Conditional expectations and operator decompositions, Ann. Global Anal. Geom. 12 (1994) 335–339.

[26] B. Simon, Trace Ideals and their Applications, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, New York, 1979.

[27] E. Vesentini, Invariant metrics on convex cones, Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser. 4 3 (1976) 671–696.