From zero Lyapunov exponents to bounded products

On the occasion of the 60th birthday of Rodrigo Bamón

Jairo Bochi

Departamento de Matemática, PUC–Rio
Credits

Joint work with Andrés Navas.

Some parts joint with Artur Avila and David Damanik.
Cocycles

- $\Omega =$ compact Hausdorff space.
- $F : \Omega \to \Omega$ continuous map.
- $G =$ topological group (main example: $\text{GL}(\mathbb{R})$).
- $A : \Omega \to G$ continuous map.

We call $A$ a **cocycle** over $F$. 

Cocycles

- $\Omega =$ compact Hausdorff space.
- $F : \Omega \to \Omega$ continuous map.
- $G = \text{topological group (main example: } \text{GL}(\cdot, \mathbb{R})\text{)}$.
- $A : \Omega \to G$ continuous map.

We call $A$ a **cocycle** over $F$.

Dynamical product (non-commutative analogs of Birkhoff sums):

$$A^{(n)}(\omega) := A(F^{n-1}\omega) \cdots A(F\omega)A(\omega).$$

Cocycle relation

$$A^{(n+m)}(\omega) = A^{(m)}(F^n\omega)A^{(n)}(\omega).$$
Skew-product

If $H$ is any space where $G$ acts, we can define a dynamical system on $\Omega \times H$:

$$T : (\omega, p) \mapsto (F(\omega), A(\omega) \cdot p).$$

So

$$F^n(\omega, x) = (F^n(\omega)), A(F^{n-1}(\omega)) \cdots A(F(\omega)))A(\omega) \cdot p).$$
Skew-product

If $H$ is any space where $G$ acts, we can define a dynamical system on $\Omega \times H$:

$$T : (\omega, p) \mapsto (F(\omega), A(\omega) \cdot p).$$

So

$$F^n(\omega, x) = (F^n(\omega)), \underbrace{A(F^{n-1}(\omega)) \cdots A(F(\omega))}_A^{(n)}(\omega)A(\omega) \cdot p).$$
Two cocycles $A$ and $B$ (over $F$) are said to be **conjugate** (or **cohomologous**) if there exists a continuous map $U : \Omega \to G$ such that

$$A(\omega) = U(F\omega)B(\omega)U(\omega)^{-1}, \quad \text{for all } \omega \in \Omega.$$ 

Then the associated skew-products on $\Omega \times H$ are conjugate under a conjugacy $(\omega, p) \mapsto (\omega, U(\omega)p)$. 

**Conjugacy**
Growth of products

Take a linear cocycle, i.e., a cocycle with $G = \text{GL}(d, \mathbb{R})$. Let $\mu$ be an $F$-invariant probability on $\Omega$. **Lyapunov exponent:**

$$L(F, A, \mu) = \lim_{n \to \infty} \frac{1}{n} \log \| A^n(\omega) \| \geq 0.$$  
(The limit exists for $\mu$-almost every $\omega$, and is independent of $\omega$.)
Growth of products

Take a linear cocycle, i.e., a cocycle with $G = \text{GL}(d, \mathbb{R})$.
Let $\mu$ be an $F$-invariant probability on $\Omega$. Lyapunov exponent:

$$L(F, A, \mu) = \lim_{n \to \infty} \frac{1}{n} \log \|A^n(\omega)\| \geq 0.$$  

(The limit exists for $\mu$-almost every $\omega$, and is independent of $\omega$.)
We say that cocycle has uniform subexponential growth if

$$\forall \varepsilon > 0 \ \exists C_\varepsilon \text{ s.t. } \|[A^n(\omega)]^{\pm 1}\| \leq C_\varepsilon e^{\varepsilon n} \ \forall \omega \in \Omega.$$  

An equivalent condition is:

$$L(F, A, \mu) = 0 \text{ for every invariant } \mu.$$
Growth of products

Take a \textit{linear} cocycle, i.e., a cocycle with $G = \text{GL}(d, \mathbb{R})$.
Let $\mu$ be an $F$-invariant probability on $\Omega$. \textbf{Lyapunov exponent:}

$$L(F, A, \mu) = \lim_{n \to \infty} \frac{1}{n} \log \|A^n(\omega)\| \geq 0.$$ 

(The limit exists for $\mu$-almost every $\omega$, and is independent of $\omega$.)
We say that cocycle has \textbf{uniform subexponential growth} if

$$\forall \epsilon > 0 \ \exists C_\epsilon \ \text{s.t.} \ \|[A^n(\omega)]^{\pm 1}\| \leq C_\epsilon e^{\epsilon n} \ \forall \omega \in \Omega.$$ 

An equivalent condition is:

$$L(F, A, \mu) = 0 \ \text{for every invariant} \ \mu.$$ 

A stronger condition: The cocycle is \textbf{product-bounded} if

$$\exists C > 0 \ \text{s.t.} \ \|[A^n(\omega)]^{\pm 1}\| \leq C \ \forall \omega \in \Omega.$$
A way get to product-bounded cocycles

Suppose $A$ is conjugate to a cocycle of “rotations” $B : \Omega \rightarrow O(d, \mathbb{R})$, i.e.,

\[ \exists \ U \text{ continuous s.t.} \]

\[ A(\omega) = U(F\omega)B(\omega)U(\omega)^{-1}. \]
A way get to product-bounded cocycles

Suppose $A$ is conjugate to a cocycle of “rotations” $B : \Omega \to O(d, \mathbb{R})$, i.e.,
exists $U$ continuous s.t.

$$A(\omega) = U(F\omega)B(\omega)U(\omega)^{-1}.$$ 

Then $A$ is obviously product-bounded, because:

$$A^{(n)}(\omega) = U(F^n\omega)B^{(n)}(\omega)U(\omega)^{-1}.$$
A way get to product-bounded cocycles

Suppose $A$ is conjugate to a cocycle of “rotations” $B : \Omega \to O(d, \mathbb{R})$, i.e.,
\[ \exists \ U \text{ continuous s.t.} \]
\[ A(\omega) = U(F\omega)B(\omega)U(\omega)^{-1}. \]

Then $A$ is obviously product-bounded, because:
\[ A^{(n)}(\omega) = U(F^n\omega)B^{(n)}(\omega)U(\omega)^{-1}. \]

Coronel–Navas–Ponce: If the dynamics $F$ is minimal then the converse holds (Product-bounded $\Rightarrow$ conjugate to rotations).
And how to get a conjugacy with rotations?

Take $H =$ space of ellipsoids in $\mathbb{R}^d$ (centered at the origin) There is a obvious action of $\text{GL}(d, \mathbb{R})$ on $H$.

(A fact to be used later: There is a Riemannian metric on $H$ which is invariant under the action, and has sectional curvature $\leq 0$.)
And how to get a conjugacy with rotations?

Take $H =$ space of ellipsoids in $\mathbb{R}^d$ (centered at the origin) There is a obvious action of $\text{GL}(d, \mathbb{R})$ on $H$.

(A fact to be used later: There is a Riemannian metric on $H$ which is invariant under the action, and has sectional curvature $\leq 0$.)

A matrix is a rotation iff it leaves invariant the unit ball $p_0 \in H$. 
And how to get a conjugacy with rotations?

Take $H =$ space of ellipsoids in $\mathbb{R}^d$ (centered at the origin) There is a obvious action of GL($d, \mathbb{R}$) on $H$.

(A fact to be used later: There is a Riemannian metric on $H$ which is invariant under the action, and has sectional curvature $\leq 0$.)

A matrix is a rotation iff it leaves invariant the unit ball $p_0 \in H$.

Then a cocycle $A$ is conjugate to a cocycle of rotations iff it has a

continuous invariant section $\phi : \Omega \to H$

Invariance: $A(\omega) \cdot \phi(\omega) = \phi(F\omega)$

Invariance means that the graph of $\phi$ is invariant under the skew-product.

Proof of the “iff”: conjugate to send $\phi$ to constant $= \text{the unit ball } \in H$. 

First result

Theorem [B.–Navas]

If $A : \Omega \to \text{GL}(d, \mathbb{R})$ has uniform subexponential growth then we can perturb it in the $C^0$-topology so that it becomes conjugate to a cocycle of rotations (and in particular, becomes product-bounded).

Done previously by Avila–B.–Damanik for the group $\text{SL}(2, \mathbb{R})$ (with some assumptions on $\Omega$).
Theorem [B.–Navas]

If \( A : \Omega \to GL(d, \mathbb{R}) \) has uniform subexponential growth then we can perturb it in the \( C^0 \)-topology so that it becomes conjugate to a cocycle of rotations (and in particular, becomes product-bounded).

Done previously by Avila–B.–Damanik for the group \( SL(2, \mathbb{R}) \) (with some assumptions on \( \Omega \)).
Applications of this kind of result

**Theorem 1 [ABD]: Genericity of Cantor spectrum**

Let $T$ be “rotation-like”. For $C^0$-generic potential functions $v : X \to \mathbb{R}$, the associated spectrum $\Sigma \subset \mathbb{R}$ is a **Cantor set**.
Applications of this kind of result

Theorem 1 [ABD]: Genericity of Cantor spectrum

Let $T$ be “rotation-like”. For $C^0$-generic potential functions $v : X \to \mathbb{R}$, the associated spectrum $\Sigma \subset \mathbb{R}$ is a Cantor set.

Theorem 2 [ABD]: Denseness of uniform hyperbolicity

Let $T$ be “rotation-like”. Then the uniformly hyperbolic matrix maps form an open and dense subset of $C^0(X, \text{SL}(2, \mathbb{R}))$.

$0$ means homotopic to constant.
Applications of this kind of result

Theorem 1 [ABD]: Genericity of Cantor spectrum

Let $T$ be “rotation-like”. For $C^0$-generic potential functions $v : X \to \mathbb{R}$, the associated spectrum $\Sigma \subset \mathbb{R}$ is a Cantor set.

Theorem 2 [ABD]: Denseness of uniform hyperbolicity

Let $T$ be “rotation-like”. Then the uniformly hyperbolic matrix maps form an open and dense subset of $C^0(\mathcal{X}, SL(2, \mathbb{R}))$.

0 means homotopic to constant.

A potential application: Try to extend Theorem 2 above for $d > 2$, replacing uniform hyperbolicity is projective hyperbolicity (also called exponential separation or domination).
Another application

A theorem of B.–Viana says that for a generic linear cocycle, the Oseledets splitting is (trivial or) dominated. If the dynamics is uniquely ergodic then we can apply the fibered versions previous results to perturb the cocycle and make the action conformal in the subbundles of the dominated splitting. More precisely, we have . . .
Another application

**Theorem**

Assume that $F : \Omega \to \Omega$ is uniquely ergodic with an invariant measure of full support. Then there is a dense subset $\mathcal{D}$ of $C^0(\Omega, \text{GL}(d, \mathbb{R}))$ such that for all $A \in \mathcal{D}$ there are:

- a Riemannian metric on the vector bundle $\Omega \times \mathbb{R}^d$ (that is, a continuous choice of inner product $\langle \cdot, \cdot \rangle_\omega$ on each fiber $\{\omega\} \times \mathbb{R}^d$);
- a continuous $(F, A)$-invariant splitting $\mathbb{R}^d = E_1(\omega) \oplus \cdots \oplus E_k(\omega)$ which is orthogonal with respect to the Riemannian metric;
- constants $c_1 > \cdots > c_k > 0$

such that, denoting $\|v\|_\omega = \sqrt{\langle v, v \rangle_\omega}$, we have

$$\|A(\omega)v_i\|_{T\omega} = c_i\|v_i\|_\omega \quad \text{for all } v_i \in E_i(\omega).$$
The simplest case

Since $\text{GL}(1, \mathbb{R}) \sim (\mathbb{R}, +)$ (take log), the case $d = 1$ of the B–N theorem reduces to:

**Baby Theorem: The one-dimensional case**

Let $g : \Omega \to \mathbb{R}$ be a continuous function. If the Birkhoff averages $\frac{1}{n} g^{(n)}$ converge uniformly to zero, then there is $\tilde{g}$ arbitrarily close to $g$ that is a coboundary.
Proof of the Baby Theorem

(Idea from [CNP2].)
Define a sequence of continuous functions $\phi_N : \Omega \to \mathbb{R}$ by the magic formula:

$$\phi_N(\omega) = \frac{1}{N} \sum_{i=0}^{N-1} [-g^{(i)}(\omega)]$$

Then $\phi_N(F\omega) = \frac{1}{N} \sum_{i=0}^{N-1} [-g^{(i+1)}(\omega) + g^{(i)}(\omega)] = \phi_N(\omega) + g(\omega) + g(N)(\omega)$.

Since $g(N)(\omega) \to 0$ (assumption), we conclude that the sequence of coboundaries $\phi_N \circ F - \phi_N$ converges uniformly to $g$. 
Proof of the Baby Theorem

(Idea from [CNP2].)

Define a sequence of continuous functions \( \phi_N : \Omega \to \mathbb{R} \) by the \textit{magic formula}:

\[
\phi_N(\omega) = \frac{1}{N} \sum_{i=0}^{N-1} [-g(i)(\omega)]
\]

Then

\[
\phi_N(F\omega) = \frac{1}{N} \sum_{i=0}^{N-1} [-g(i)(F\omega)] = \frac{1}{N} \sum_{i=0}^{N-1} [-g(i+1)(\omega) + g(\omega)]
\]

\[
= \phi_N(\omega) + g(\omega) + \frac{g^{(N)}(\omega)}{N}.
\]

Since \( \frac{g^{(N)}(\omega)}{N} \to 0 \) (assumption), we conclude that the sequence of coboundaries \( \phi_N \circ F - \phi_N \) converges uniformly to \( g \).
Proof of the Adult Theorem: ideas

We need to perturb $A$ to $\tilde{A}$ so that the associated skew-product has an invariant section of ellipsoids. Two steps:

1. First show that if $A$ has uniform subexponential growth, then there is a sequence of "almost invariant" sections $\phi_N: \Omega \to H$; more precisely, the distance between the graph of $\phi_N$ and its $F$-image tends to zero as $N \to \infty$.

2. Next we will show that if $(F, A)$ has a almost invariant section $\phi$, then one can perturb $A$ so that $\phi$ becomes invariant.
Proof of the Adult Theorem: ideas

We need to perturb $A$ to $\tilde{A}$ so that the associated skew-product has an invariant section of ellipsoids. Two steps:

1. First show that if $A$ has uniform subexponential growth, then there is a sequence of “almost invariant” sections $\phi_N : \Omega \to H$; more precisely, the distance between the graph of $\phi_N$ and its $F$-image tends to zero as $N \to \infty$. 

Jairo Bochi (PUC-Rio)  
From zero Lyapunov exponents to bounded products  
November 23, 2011 15 / 25
Proof of the Adult Theorem: ideas

We need to perturb $A$ to $\tilde{A}$ so that the associated skew-product has an invariant section of ellipsoids. Two steps:

1. First show that if $A$ has uniform subexponential growth, then there is a sequence of “almost invariant” sections $\phi_N : \Omega \to H$; more precisely, the distance between the graph of $\phi_N$ and its $F$-image tends to zero as $N \to \infty$.

2. Next we will show that if $(F, A)$ has an almost invariant section $\phi$, then one can perturb $A$ so that $\phi$ becomes invariant.
Clues on how to perform each of the steps

1. For Step 1, the idea to construct almost invariant sections $\phi_N$ is to imitate the magic formula, replacing the averaging operation $\frac{1}{N} \sum_{i=0}^{N-1}$ by an appropriate barycenter concept.
Clues on how to perform each of the steps

1. For Step 1, the idea to construct almost invariant sections $\phi_N$ is to imitate the magic formula, replacing the averaging operation
$$\frac{1}{N} \sum_{i=0}^{N-1} \text{ by an appropriate } \textit{barycenter} \text{ concept.}$$

2. In Step 2, the perturbation $\tilde{A}(\omega)$ of $A(\omega)$ is obtained by left-multiplying $A(\omega)$ by a matrix close to the identity that takes the ellipsoid $A(\omega)\phi(\omega)$ to the nearby ellipsoid $\phi(F\omega)$. 

(Rem: Both concepts were introduced by Cartan in the 1920's.)
Clues on how to perform each of the steps

1. For Step 1, the idea to construct almost invariant sections $\phi_N$ is to imitate the magic formula, replacing the averaging operation $\frac{1}{N} \sum_{i=0}^{N-1}$ by an appropriate barycenter concept.

2. In Step 2, the perturbation $\tilde{A}(\omega)$ of $A(\omega)$ is obtained by left-multiplying $A(\omega)$ by a matrix close to the identity that takes the ellipsoid $A(\omega)\phi(\omega)$ to the nearby ellipsoid $\phi(F\omega)$. This is not as trivial as it seems, because these two ellipsoids may be very distorted, and moreover the correcting matrices must be chosen in a way that depends continuously on $\omega$. We accomplish this by establishing a certain uniform homogeneity property of the space of ellipsoids.
Clues on how to perform each of the steps

1. For Step 1, the idea to construct almost invariant sections $\phi_N$ is to imitate the magic formula, replacing the averaging operation $\frac{1}{N} \sum_{i=0}^{N-1}$ by an appropriate barycenter concept.

2. In Step 2, the perturbation $\tilde{A}(\omega)$ of $A(\omega)$ is obtained by left-multiplying $A(\omega)$ by a matrix close to the identity that takes the ellipsoid $A(\omega)\phi(\omega)$ to the nearby ellipsoid $\phi(F\omega)$. This is not as trivial as it seems, because these two ellipsoids may be very distorted, and moreover the correcting matrices must be chosen in a way that depends continuously on $\omega$. We accomplish this by establishing a certain uniform homogeneity property of the space of ellipsoids.

The key tools, the barycenter and the symmetries, are geometric and work in more general situations. . .
(Rem: Both concepts were introduced by Cartan in the 1920’s.)
Cocycles of isometries

\( H = \) symmetric simply-connected space of curvature \( \leq 0 \). (Possibly of infinite dimension.)

We consider **cocycles of isometries**, i.e. cocycles in the group \( \text{Isom}(H) \).

Example: \( H = \) Poincaré disc.
A familiar example of $H$

Let $\mathbb{H} = \{x + iy \in \mathbb{C}; \ y > 0\}$ be the hyperbolic half-space; Riemannian metric $ds^2 = (dx^2 + dy^2)/y^2$.

The group $\text{SL}(2, \mathbb{R})$ acts on $\mathbb{H}$ by isometries:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \phi_A(z) = \frac{az + b}{cz + d}$$
A familiar example of $H$

Let $H = \{x + iy \in \mathbb{C}; \ y > 0\}$ be the hyperbolic half-space; Riemannian metric $ds^2 = (dx^2 + dy^2)/y^2$.

The group $SL(2, \mathbb{R})$ acts on $H$ by isometries:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \phi_A(z) = \frac{az + b}{cz + d}$$

$SL(2, \mathbb{R})$ also acts on the hyperbolic disk $D = \{z \in \mathbb{C}; \ |z| < 1\}$: conjugate the previous action with a Möbius map that takes $H$ to $D$.

The dynamics of an isometry ($\neq \text{id}$) looks like one of these:
Back to cocycles of isometries: $F : \Omega \to \Omega$, $A : \Omega \to \text{Isom}(H)$.

Analogue of the (upper) Lyapunov exponent:

$$\text{drift}(F, A) = \lim_{n \to \infty} \frac{1}{n} \text{dist}(A^n(x) \cdot p_0, p_0)$$

(a measure of the speed the orbits approach $\partial H$.)

∃ generalization of Oseledets Theorem to this context (Kaimanovich, Karlsson, Margulis, Ledrappier).
The main result

Theorem [B.–Navas]

Given a cocycle of isometries $A : \Omega \to \text{Isom}(H)$ (over a dynamics $T : \Omega \to \Omega$) with sublinear drift to infinity, we can perturb it to create an invariant section $\phi : \Omega \to H$. In particular, the perturbed cocycle has bounded orbits.
The main result

Theorem [B.–Navas]

Given a cocycle of isometries $A : \Omega \to \text{Isom}(H)$ (over a dynamics $T : \Omega \to \Omega$) with sublinear drift to infinity, we can perturb it to create an invariant section $\phi : \Omega \to H$. In particular, the perturbed cocycle has bounded orbits.

Remarks:

1. ∃ similar theorem for continuous–time dynamical systems.
The main result

Theorem [B.–Navas]

Given a cocycle of isometries $A : \Omega \rightarrow \text{Isom}(H)$ (over a dynamics $T : \Omega \rightarrow \Omega$) with sublinear drift to infinity, we can perturb it to create an invariant section $\phi : \Omega \rightarrow H$. In particular, the perturbed cocycle has bounded orbits.

Remarks:

1. ∃ similar theorem for continuous–time dynamical systems.
2. ∃ a non-perturbative result on the existence of section of nearly minimal displacement. The hypotheses are weaker (Buseman, CAT(0), . . . )
The main result

Theorem [B.–Navas]

Given a cocycle of isometries $A : \Omega \to \text{Isom}(H)$ (over a dynamics $T : \Omega \to \Omega$) with sublinear drift to infinity, we can perturb it to create an invariant section $\phi : \Omega \to H$. In particular, the perturbed cocycle has bounded orbits.

Remarks:

1. ∃ similar theorem for continuous–time dynamical systems.
2. ∃ a non-perturbative result on the existence of section of nearly minimal displacement. The hypotheses are weaker (Buseman, CAT(0), . . .)
3. ∃ versions of these result for other fiber bundles ($\neq \Omega \times H$). (Those are actually needed in our application.)
First tool: barycenter

The Cartan barycenter of a list of points $p_1, \ldots, p_n$ is the unique point that minimizes the function:

$$x \mapsto \sum_{i=1}^{n} d(x, p_i)^2.$$  

(Works globally with our assumptions on $H$)

The barycenter is obviously equivariant under isometries.
First tool: barycenter

The Cartan barycenter of a list of points $p_1, \ldots, p_n$ is the unique point that minimizes the function:

$$x \mapsto \sum_{i=1}^{n} d(x, p_i)^2.$$  

(Works globally with our assumptions on $H$)

The barycenter is obviously equivariant under isometries.

“Lipschitzness” property:

$$d(\text{bar}\{p_i\}, \text{bar}\{q_i\}) \leq \frac{1}{n} \sum_{i=1}^{n} d(p_i, q_i).$$
Step 1: The magic formula for the almost-invariant section

\[
\phi_N(x) = \text{bar} \left( p_0, A(x)^{-1} \cdot p_0, [A^{(2)}(x)]^{-1} \cdot p_0, \ldots, [A^{(N-1)}(x)]^{-1} \cdot p_0 \right)
\]
Uniform homogeneity of $H$

To perform Step 2 of the proof of the Main Theorem (the closing of the almost-invariant section), we use the following:

**Lemma (Macroscopic uniform homogeneity)**

There exists a continuous map $J : H \times H \to \text{Isom}(H)$ with the following properties:

1. $J(p, q)p = q$ for all $p, q \in H$.
2. $J(p, q)$ converges **uniformly** on bounded sets of $H$ to the identity as the distance between $p$ and $q$ converges to zero.

(Notice that $p$ and $q$ are not restricted to a bounded set; that is essentially what makes the lemma non-trivial.)
Uniform homogeneity of $H$

To perform Step 2 of the proof of the Main Theorem (the closing of the almost-invariant section), we use the following:

**Lemma (Macroscopic uniform homogeneity)**

There exists a continuous map $J : H \times H \to \text{Isom}(H)$ with the following properties:

1. $J(p, q)p = q$ for all $p, q \in H$.
2. $J(p, q)$ converges uniformly on bounded sets of $H$ to the identity as the distance between $p$ and $q$ converges to zero.

More explicitly, assertion 2 means:

$$\forall \varepsilon > 0 \; \forall B \subset H \text{ bounded} \; \exists \delta > 0 \text{ such that}$$

$$p, q \in H, \; d(p, q) < \delta \Rightarrow d(J(p, q)r, r) < \varepsilon \; \forall r \in B.$$

(Notice that $p$ and $q$ are not restricted to a bounded set; that is essentially what makes the lemma non-trivial.)
Proof of the uniform homogeneity lemma

For any \( p \in H \), let \( \sigma_p : H \to H \) denote the Cartan symmetry around \( p \). We want to continuously define a isometry \( J(p, q) \) that takes \( p \) to \( q \) and moves the base-point \( p_0 \) as little as possible.
Proof of the uniform homogeneity lemma

For any $p \in H$, let $\sigma_p : H \to H$ denote the Cartan symmetry around $p$. We want to continuously define a isometry $J(p, q)$ that takes $p$ to $q$ and moves the base-point $p_0$ as little as possible.

$m$ is the midpoint of $q$ and $\sigma_{p_0}(p)$. The isometry $J(p, q) := \sigma_m \circ \sigma_{p_0}$ sends $p$ to $q$ and translates the geodesic joining $p_0$ and $m$ by length $2\text{dist}(p_0, m)$, which by nonpositive curvature is $\leq \text{dist}(p, q)$. 
Thank you!

Happy birthday, Rodrigo!