

The Minimal Entropy Theorem and Mostow Rigidity

after G. Besson, G. Courtois and S. Gallot

1 Introduction

A classical “rigidity” result in Geometry, which is similar in form to the “entropy rigidity” theorem to be discussed in these notes is the well known *isoperimetric inequality*: Let D be a bounded domain in \mathbb{R}^2 with smooth boundary, let A be the area of D and L the perimeter of its boundary. Then, the geometric invariant L^2/A satisfies the inequality

$$L^2/A \geq 4\pi,$$

and equality holds exactly when D is a disk.

More generally, given a smooth manifold M , one may have on a space $\mathcal{G}(M)$ of Riemannian metrics on M a functional

$$\mathcal{F} : \mathcal{G}(M) \rightarrow \mathbb{R},$$

satisfying some inequality $\mathcal{F}(g) \geq c_0$, and the goal is to characterize the minimizing metrics.

The functional \mathcal{F} considered below is associated to the *volume entropy* of Riemannian manifolds with negative sectional curvature. It is an asymptotic measure of growth for the volumes of balls in the universal cover of M .

More precisely, let M be a compact, connected, smooth n -manifold, and g a Riemannian metric on M . Let X be the universal covering of M and denote by $V_p^g(R) := \text{Vol}(B_p^g(R))$ the volume of the ball of radius R centered at the point $p \in X$. The (volume) *entropy* of g is defined as

$$h(g) := \lim_{R \rightarrow +\infty} \frac{1}{R} \log V_p^g(R).$$

(For the relationship of $h(g)$ with other notions of entropy and its dynamical and probabilistic significance see, for example, [9].)

Some elementary (and easy) facts about $h(g)$ are stated in the next proposition.

Proposition 1.1 *On a compact, smooth n -manifold M the limit that defines $h(g)$ always exists and is independent of the choice of point p in the universal covering of M . (See [11].) Moreover, the functional*

$$\mathcal{F}(g) := [h(g)]^n \text{Vol}(M, g)$$

has the property that $\mathcal{F}(g') = \mathcal{F}(g)$, for any two homothetic metrics; that is, such that $g' = \lambda g$ for some $\lambda > 0$. If M has constant sectional curvature $K \equiv -\kappa^2$, then

$$h(g) = (n-1)\kappa.$$

In fact, in this case

$$V_p^g(R) = c_n \int_0^R (\kappa^{-1} \sinh(\kappa\rho))^{n-1} d\rho \sim c' e^{(n-1)\kappa R}.$$

Note, therefore, that \mathcal{F} is a dimensionless quantity, which remains the same whether distances in M are measured in inches, centimeters or light-years. The next (main) theorem is proved in [4]. It refers to the functional \mathcal{F} just defined.

Theorem 1.2 (Besson-Courtois-Gallot) *Let M and N be two compact, connected, oriented smooth manifolds of same dimension n . Suppose that there exists a continuous function $f : M \rightarrow N$ with nonzero degree and that N is equipped with a locally symmetric Riemannian metric g_0 of negative curvature. Then, for any metric g on M , one has*

$$\mathcal{F}(g) \geq |\deg(f)| \mathcal{F}(g_0).$$

Moreover, if $n \geq 3$, equality is achieved exactly when (M, g) is locally symmetric and there exists a positive constant λ such that $(M, \lambda g)$ is a Riemannian covering of (N, g_0) , with covering map homotopic to f .

For the many consequences of the theorem, we refer the reader to [4]. The theorem gives a complete answer to a question first posed in a more dynamical context by A. Katok in [8] and later posed in the more Riemannian geometric setting by M. Gromov (see [6, 7]). We only note the following immediate corollary.

Corollary 1.3 (Mostow) *Suppose that M and N are two compact, locally symmetric spaces of dimension n , $n \geq 3$, with strictly negative sectional curvature. Suppose that M and N are homotopically equivalent. (This last condition is equivalent to their fundamental groups being isomorphic.) Then, if $\text{Vol}(M) = \text{Vol}(N)$, the two manifolds are isometric.*

Proof. Since M and N are homotopically equivalent, there exists $f : M \rightarrow N$ of degree 1. Moreover, since now both spaces are locally symmetric, the inequality of the theorem becomes an equality. Therefore, the two manifolds are homothetic. As they have the same volume they are, in fact, isometric. \square

We recall that a locally symmetric space of negative curvature is locally isometric (up to homothety) to one of the hyperbolic spaces over the reals, complex numbers, quaternions, or the Cayley number. (The last case only occurs in real dimension 16, which corresponds to dimension 2 over the Cayley numbers.)

In these notes we shall prove a special case of the theorem, namely, we shall assume that (N, g_0) has constant negative curvature (equivalently, it is locally isometric to the real hyperbolic space) and that (M, g) has (*a priori* variable) negative curvature. Moreover, $f : M \rightarrow N$ will be assumed to be a homotopy equivalence. The first assumption is made for the sake of notational simplicity, but the condition that the curvature of (M, g) is negative (as well as the last extra condition) introduces genuine simplifications. (See [4].)

It is well known that Mostow's theorem does not hold in dimension 2. In fact, the space of nonisometric (but locally isometric) metrics of constant curvature -1 on a compact orientable surface S is (essentially) the Teichmüller space, which is an analytic manifold of real dimension $6[\text{genus}(S) - 1]$. It is interesting to isolate what goes wrong in the proof that we shall discuss. What breaks down is the claim of the the following exercise.

Proposition 1.4 (The brain in a jar.) *Let H be an $n \times n$ positive symmetric matrix, with $\text{trace}(H) = 1$. Assume that $n \geq 3$. Then*

$$\frac{\det(H)}{[\det(I - H)]^2} \leq \left[\frac{n}{(n-1)^2} \right]^n$$

and equality holds if and only if

$$H = \frac{1}{n}I.$$

Moreover, the same is not true for $n = 2$.

(Hint: Use Lagrange multipliers.)

2 Patterson-Sullivan measures

The construction of the homothety claimed in the main theorem will involve a certain imbedding of the universal covering of M into the space of probability measures of its ideal boundary. To make sense of this claim, we need to introduce a compactification of the universal covering, obtained by attaching a boundary at infinity.

Let X be a proper metric space, i.e. closed balls are compact. The group $\text{Is}(X)$ of isometries of X endowed with the compact-open topology is locally compact and second countable. It acts on X with closed orbits and compact stabilizers. Moreover, for each $p \in X$, denote by $d_p(q) = d(p, q)$ the distance to p . Denote by $C(X)$ the space of continuous functions on X with the topology of uniform convergence on compact sets. $C(X)$ contains \mathbb{R} as the one-dimensional subspace of constant functions. Define $C_*(X) = C(X)/\mathbb{R}$, the space of equivalence classes of continuous functions, where two functions $f, g \in C(X)$ are equivalent if and only if they differ by a constant. Convergence in the quotient topology can be described as follows: α_i converges to α in $C_*(X)$ if and only if there are representatives f_i, f in $C(X)$ such that $\alpha_i = [f_i]$, $\alpha = [f]$ and f_i converges to f in X . The proof of the following proposition is straightforward.

Proposition 2.1 *The map*

$$i : p \in X \mapsto [d_p] \in C_*(X)$$

is an embedding and $C_(X)$ is homeomorphic with the subspace of $C(X)$ consisting of functions which vanish at a fixed point p_0 . Show that $C(X)/\mathbb{R}$ is Hausdorff.*

Definition 2.2 *Define the closure $\text{Cl}(X)$ of X as the closure of $i(X)$. The ideal boundary of X , denoted $X(\infty)$ is the complement of $i(X)$ in $\text{Cl}(X)$. We refer to the elements of $X(\infty)$ as the points at infinity of X .*

It is immediate from the definitions that for any $\alpha = [h] \in \text{Cl}(X)$ and points $p, q \in X$, the difference $h(p) - h(q)$ is well defined (independently of the representative h) and

$$|h(p) - h(q)| \leq d(p, q).$$

It follows from this equicontinuity property and Arzela-Ascoli Theorem that $\text{Cl}(X)$ is compact.

Definition 2.3 *A continuous function $h \in C(X)$ such that $[h] \in X(\infty)$ is called a horofunction. The sublevel sets $h^{-1}(-\infty, c)$ are called (open) horoballs and the level sets $h^{-1}(c)$ are the horospheres. A horosphere associated to a point at infinity can thus be regarded as the set of points on X “equidistant” to that point at infinity. If $\alpha = [h] \in X(\infty)$ and $p, q \in X$, we denote*

$$B_\alpha(p, q) := h(p) - h(q).$$

The construction of the Patterson-Sullivan measures that we describe next is taken from [5]. We continue to assume that X is an arbitrary complete Riemannian manifold.

Definition 2.4 *For a positive Radon measure m on X , the number*

$$\delta := \inf\{s \in [0, \infty] : \int_X e^{-sd(p, q)} dm(q) < +\infty\}$$

is independent of $p \in X$ and is called the critical exponent of m . The critical exponent δ_G of a closed subgroup G of the group $\text{Is}(X)$ of isometries of X is by definition the critical exponent of a positive G -invariant measure supported on a G -orbit in X . Equivalently,

$$\delta_G := \inf\{s \in [0, \infty] : \int_G e^{-sd(p, g(q))} dg < +\infty\}.$$

We denote by $\mathcal{M}^+(X(\infty))$ the space of positive Radon measures on the ideal boundary of X . Then an α -dimensional density for a closed subgroup G in $\text{Is}(X)$ is a continuous G -equivariant map

$$\mu : p \in X \mapsto \mu_p \in \mathcal{M}^+(X(\infty))$$

such that

$$\frac{d\mu_p}{d\mu_q}(\xi) = e^{-\alpha B_\xi(p, q)},$$

for all $p, q \in X$. (G -equivariance means that for all $p \in X$ and $g \in G$, $g_\mu_p = \mu_{g(p)}.$)*

Proposition 2.5 *Let m be a positive Radon measure on X , δ its critical exponent, $G := \{g \in \text{Is}(X) | g_* m = m\}$. Assume $m(X) = +\infty$ and $\delta < +\infty$. Then there exists a δ -dimensional density $p \mapsto \mu_p$ for G such that*

$$\text{supp}(\mu_p) \subset \overline{\text{supp}(m)} \cap X(\infty)$$

for all $p \in X$.

The construction depends on the following simple lemma due to Patterson [13], concerning measures on the real line.

Lemma 2.6 *Let μ be a positive Radon measure on $\mathbb{R}_+ := [0, \infty)$. Define the exponent of μ as*

$$\delta := \delta(\mu) := \inf\{s \in \mathbb{R}_+ | \int_0^\infty e^{-st} d\mu(t) < +\infty\}.$$

Assume that δ is finite and positive. Show that there exists a continuous increasing function $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that

1. *the integral $\int_0^\infty e^{-st} h(e^t) d\mu(t)$ converges for $s > \delta$ and diverges at $s = \delta$;*
2. *for all $\epsilon > 0$, there exists t_0 such that for all $t \geq t_0$ and $s > 1$*

$$h(st) \leq t^\epsilon h(s).$$

Proof. Choose a decreasing sequence $\epsilon_n > 0$ approaching 0. We define a sequence t_1, t_2, \dots increasing to ∞ and h as follows. Set $t_1 = 1$ and $h|_{[0, t_1]} \equiv 1$. Suppose t_n and $h|_{[0, t_n]}$ already defined. Choose t_{n+1} such that

$$\frac{h(t_n)}{t_n^{\epsilon_n}} \int_{\ln t_n}^{\ln t_{n+1}} e^{-(\delta - \epsilon_n)t} d\mu(t) \geq 1.$$

For $t \in [t_n, t_{n+1}]$ define

$$h(t) := \frac{h(t_n)}{t_n^{\epsilon_n}} t^{\epsilon_n}.$$

We claim that the integral $\int_0^\infty e^{-\delta t} h(e^t) d\mu(t)$ diverges. In fact,

$$\int_1^\infty e^{-\delta t} h(e^t) d\mu(t) = \sum_{n=1}^\infty \int_{\ln t_n}^{\ln t_{n+1}} e^{-\delta t} h(e^t) d\mu(t)$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} \int_{t_n}^{t_{n+1}} u^{-\delta} \frac{h(u)}{u} d\mu(u) \\
&= \sum_{n=1}^{\infty} \int_{t_n}^{t_{n+1}} u^{-(\delta-\epsilon_n)} \frac{h(t_n)}{t_n^{\epsilon_n}} \frac{1}{u} d\mu(u) \\
&= \sum_{n=1}^{\infty} \frac{h(t_n)}{t_n^{\epsilon_n}} \int_{t_n}^{t_{n+1}} u^{-(\delta-\epsilon_n)} \frac{1}{u} d\mu(u) \\
&= \sum_{n=1}^{\infty} \frac{h(t_n)}{t_n^{\epsilon_n}} \int_{\ln t_n}^{\ln t_{n+1}} e^{-(\delta-\epsilon_n)t} \frac{1}{u} d\mu(t) \\
&\geq \sum_{n=1}^{\infty} 1 \\
&= \infty.
\end{aligned}$$

The second part of the proposition can be seen as follows. From the definition of h we have

$$\ln h(t) - \ln h(t_n) = \epsilon_n \ln \frac{t}{t_n}.$$

It follows that $\ln h(t)$ is a piecewise continuous function of $\ln t$ with positive slope bounded above by ϵ_n on each interval $[t_n, t_{n+1}]$. For any positive ϵ , choose n such that $\epsilon > \epsilon_n$. Then for all $t > 1$ and all $s > t_n$,

$$\ln h(st) - \ln h(t) \leq \epsilon(\ln st - \ln t),$$

from which the claim follows. \square

We start now the proof of the proposition. Given a measure m on X satisfying the hypothesis of the proposition, define for $p, p_0 \in X$, $s > \delta$ and $\zeta \in \text{Cl}(X)$:

$$d\mu_{s,p}(\zeta) := \frac{e^{-sd(p,\zeta)} h(e^{d(p,\zeta)}) dm(\zeta)}{\int_X e^{-sd(p_0,q)} h(e^{d(p_0,q)}) dm(q)},$$

where $h \equiv 1$ if $\delta = 0$ and, if $\delta > 0$, h is the function associated by the lemma to the direct image of m via the proper map $\zeta \in X \mapsto d(p_0, \zeta) \in [0, \infty)$. (One easily checks the equivariance property $g_* \mu_{s,p} = \mu_{s,g(p)}$.) Using the properties of h , one verifies that the family of G -equivariant continuous maps $p \in X \mapsto \mu_{s,p} \in \mathcal{M}^+(\text{Cl}(X))$, for $\delta < s \leq \delta + 1$ is equicontinuous and uniformly bounded on compact sets. This family is therefore relatively compact in the space of continuous maps $C(X, \mathcal{M}^+(X(\infty)))$ endowed with the topology of uniform convergence on compact sets. It follows from part (1) of the lemma that any accumulation point $x \mapsto \mu_x$ of this family takes its values in $\mathcal{M}^+(X(\infty))$.

One has for $p, q \in X$ and $s > \delta$:

$$\frac{d\mu_{s,p}}{d\mu_{s,q}}(\zeta) = e^{-sB_\zeta(p,q)} \frac{h(d(p,\zeta))}{h(d(q,\zeta))}.$$

The function

$$\zeta \mapsto \frac{h(d(p,\zeta))}{h(d(q,\zeta))}$$

extends continuously to the closure of X and its value on the ideal boundary $X(\infty)$ is 1, due to part (2) of the lemma. The family of functions

$$\zeta \in \text{Cl}(X) \mapsto e^{-sB_\zeta(p,q)} \in \mathbb{R},$$

for $\delta \leq s \leq \delta + 1$ is compact. It follows that any accumulation point $p \mapsto \mu_p$ of the family $\{p \mapsto \mu_{s,p} | \delta \leq s \leq \delta + 1\}$ verifies

$$\frac{d\mu_p}{d\mu_q}(\zeta) = e^{-sB_\zeta(p,q)}$$

and is, therefore, a δ -dimensional density for G . This concludes the proof of the proposition.

3 Convexity and the barycenter

A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is *convex* if for every $a < b$ and $s \in (0, 1)$

$$f(a + s(b - a)) \leq f(a) + s(f(b) - f(a)).$$

The function is *strictly* convex if the above inequality is strict. A subset W of a Riemannian manifold M is *convex* if for all $p, q \in W$, there exists a unique shortest geodesic from p to q (disregarding the parametrization) contained in W . A function $g : M \rightarrow \mathbb{R}$ is (strictly) convex if for every nontrivial geodesic $c : [0, 1] \rightarrow M$, $g \circ c$ is (strictly) convex.

Proposition 3.1 *Let f be a differentiable convex function defined on a convex subset W of a Riemannian manifold M . Then the only critical points of f in the interior of W are the absolute minima. If f is strictly convex, there can be at most one critical point.*

Proposition 3.2 *Let M be a Riemannian manifold of (negative) nonpositive sectional curvature. Then the distance function $d : W \times W \rightarrow [0, \infty)$ is (strictly) convex for every convex subset W of M .*

Proof. This means that the function

$$t \mapsto d(c_1(t), c_2(t))$$

is convex whenever c_1 and c_2 are geodesics in W . The metric properties of the distance function d , such as convexity, and sectional curvature are related through the study of the norm of Jacobi fields. (See [10, 3.8.1, page 350].) \square

Theorem 3.3 (Hadamard-Cartan) *Let M be a complete Riemannian manifold with nonpositive sectional curvature. Then for any $p \in M$, the exponential map $\exp_p : T_p M \rightarrow M$ is a covering map. In particular, the universal covering of M is diffeomorphic to \mathbb{R}^n and M is a $K(\pi, 1)$ -space. (The homotopy type is completely determined by the fundamental group.)*

The ideal boundary of the universal covering X of a Riemannian manifold M of nonpositive curvature can be characterized in a very concrete way as follows. We say that two unit speed geodesics $c_1, c_2 : \mathbb{R} \rightarrow X$ are *asymptotic*, $c_1 \sim c_2$, if there exists a positive constant a such that $d(c_1(t), c_2(t)) \leq a$, for all $t \geq 0$. One obtains this way an equivalence relation on the set of geodesic rays. The equivalence classes will be denoted $[c]$ and the set of equivalence classes, \mathcal{R} .

Given $z_1, z_2 \in \mathcal{R}$ and $p \in M$, there exist (unique) geodesic rays c_1 and c_2 from p , with unit speed, such that $z_i = [c_i]$. Define

$$\angle_p(z_1, z_2) := \text{angle between } c'_1(0) \text{ and } c'_2(0).$$

Define similarly the angle between $z \in \mathcal{R}$ and $q \in X$. Also define the cone

$$\mathcal{C}_p(z, \epsilon) := \{q \in X \cup \mathcal{R} \mid p \neq q, \angle_p(z, q) < \epsilon\}.$$

The *cone topology* on $X \cup \mathcal{R}$ is the topology generated by open sets in X and these cones. The induced topology on \mathcal{R} is called the *sphere topology*. We have that q_i converges to $z \in \mathcal{R}$ if and only if for every $p \in X$, $d(p, q_i) \rightarrow \infty$ and $\angle_p(z, q_i) \rightarrow 0$.

There exists a homeomorphism between the unit sphere $S^{n-1} \subset T_p M$ and \mathcal{R} which associates to each unit vector v at p the class $[c_v]$ represented by the geodesic ray issuing from p with initial velocity v . One shows that $X \cup \mathcal{R}$ becomes a compactification of X which is homeomorphic to the unit closed ball in \mathbb{R}^n .

An isometry γ of X induces a homeomorphism of the *sphere at infinity* \mathcal{R} .

Proposition 3.4 *The two definitions of boundary, $X(\infty)$ and \mathcal{R} are equivalent. More precisely, there is a homeomorphism between $X(\infty)$ and \mathcal{R} which conjugates the induced action of the group of isometries.*

Proof. The correspondence is given by the map which associates to each class $[c]$ of geodesic rays the horofunction

$$h_{[c]}(p) := \lim_{t \rightarrow \infty} [d_{c(t)}] = \lim_{t \rightarrow \infty} [d(x, c(t)) - t].$$

□

The function $h_{[c]} \in C(X)/\mathbb{R}$ defined above is called a Busemann function for c .

Proposition 3.5 *Let $h \in C(X)$. Then the following are equivalent.*

1. $[h]$ is a Busemann function;
2. $[h]$ is a horofunction;
3. h satisfies the following conditions:
 - (a) h is convex,
 - (b) $|h(p) - h(q)| \leq d(p, q)$ for all $p, q \in X$
 - (c) for all $p \in X$, $r > 0$, there are $q_1, q_2 \in \partial B_r(p)$ such that

$$|h(q_1) - h(q_2)| = 2r$$

4. h is a convex C^1 function with $\|\text{grad } h\| \equiv 1$.

As an example we consider the case $K = -1$ and show how one can describe the second fundamental form of horospheres in $\mathbb{R}H^n$. First recall that given a hypersurface N of a smooth Riemannian manifold X , and ν a smooth local section of the normal bundle of N , one defines the *second fundamental form* of N as the bilinear form on TN such that for $u, v \in T_p N$,

$$L_p(u, v) = -\langle \nabla_u \nu, v \rangle.$$

One easily shows that L is symmetric.

Let h be a C^1 function on X such that its exterior derivative dh has norm $\|dh\| \equiv 1$. Equivalently, if $\nu = \text{grad } h$, assume that $\|\nu\| \equiv 1$. Define the Hessian H_h of h as the bilinear form on TN given by

$$H_p(h)(u, v) = (\nabla_u dh)(v) \text{ for } u, v \in T_p N.$$

One checks that the Hessian of h is the (negative of the) second fundamental of the level hypersurfaces of h .

Proposition 3.6 *Let M be a compact Riemannian manifold of constant sectional curvature $K \equiv -1$ and X its universal covering. Denote by g_0 the Riemannian metric on X . Let $h = h_\theta$ be a Busemann function on X , where θ is an element of the ideal boundary of X . Then, for all $p \in X$ and $u, v \in T_p X$,*

$$H_p(h)(u, v) = g_0(u, v) - dh(u)dh(v).$$

Proof. Let ν be the gradient vector field of h . Define the function

$$c : X \times \mathbb{R} \rightarrow X$$

so that $c_t(p) = c(p, t)$ is the geodesic curve through p with initial vector $\nu(p)$. Given $v \in \nu(p)^\perp$, define $Y(t) := d(c_t)_p v$. Then $Y(t)$ is a Jacobi field along the geodesic line $c_t(p)$ such that $Y(0) = v$ and

$$\lim_{t \rightarrow -\infty} \|Y(t)\| = 0$$

since for any two points p and q in the same level hypersurface of h ,

$$\lim_{t \rightarrow -\infty} d(c_t(p), c_t(q)) = 0.$$

Moreover, a simple computation shows that $\frac{\nabla Y}{dt}(t) = \nabla_\nu \nu$ so that

$$(\nabla_\nu dh)(v) = \frac{d}{dt} \|Y(t)\|_{t=0} > 0, \text{ for } v \neq 0.$$

This shows that if the sectional curvature of M is negative, the Hessian of h is positive definite on $\nu(p)^\perp$. If $K \equiv -1$, an elementary calculation shows that $Y(t) = e^t v(t)$, where $v(t)$ is the parallel translation of v along $c_t(p)$. Therefore, $H_p(h)(v, v) = g_0(v, v)$, and the claim for general u, v follows. \square

Fix a point $p_0 \in X$ and consider for each $\theta \in X(\infty)$ the unique Busemann function h_θ of θ such that $h_\theta(p_0) = 0$. Let μ be a probability measure on the compact space $X(\infty)$ and define the average

$$h_\mu(p) := \int_{X(\infty)} h_\theta(x) d\mu(\theta).$$

(Choosing a different p_0 changes h_μ by a constant.) The unique critical point of the next proposition is the *barycenter* of μ .

Proposition 3.7 *If μ has no atoms, h_μ has a unique critical point in X , corresponding to the (unique) absolute minimum.*

Proof. We first show that h_μ is strictly convex. It suffices for that to show that the quadratic form ∇dh_μ is positive definite. (Notice that, given any geodesic curve $c(t)$,

$$\frac{d}{dt}(h_\mu \circ c)|_{t=0} = \frac{\nabla}{dt}(dh_\mu(c'(t))) = (\nabla dh_\mu)(c'(0), c'(0))$$

and convexity is implied by the positivity of the second derivative in t .) Now, for each $x \in X$ and $u \in T_x X$, we have seen that $(\nabla dh_\theta)_x(u, u) > 0$ for all $\theta \neq \theta_0^+, \theta_0^-$, where θ_0^\pm are the points at infinity represented by geodesic rays that issue from x , along the directions $\pm u$. As $\mu(\{\theta_0^+, \theta_0^-\}) = 0$, we have the claim.

Consider now the set

$$A_c := \{p \in X | h_\mu(p) \leq c\}.$$

A_c is a convex set since h_μ is convex, and for $c \geq 0$, $p_0 \in A_c$. Moreover, $X = \cup_{c>0} \text{int}(A_c)$ and each set $\text{int}(A_c)$ has at most one critical point, corresponding to a minimum. Note that each $p \in A_c$ can therefore be joined to p_0 by a geodesic ray in A_c . Therefore the proposition will be proved if we show that for any geodesic ray $c(t)$ issuing from p along an arbitrary direction u ,

$$\lim_{t \rightarrow \infty} h_\mu(c(t)) = +\infty.$$

Remark that, as h_θ is convex,

$$h_\theta(c(0 + sd(p, p_0))) \leq h_\theta(c(0)) + s[h_\theta(d(p, p_0)) - h_\theta(p_0)],$$

so that ($h_\theta(p_0) = 0$)

$$h_\theta(q) \leq \frac{d(q, p_0)}{d(p, p_0)} h_\theta(p).$$

For each p on the ray $p_0\theta$, let

$$J_{\theta_0}(p) = \{\theta \in X(\infty) | h_\theta(x) \leq 0\}.$$

The function $h_\theta(x)$ is continuous in θ , so that for each p , $J_{\theta_0}(p)$ is a compact subset of $X(\infty)$. For every q on the ray from p_0 to θ_0 , situated between p and θ_0 , we have due to the above inequality that $J_{\theta_0}(q) \subset J_{\theta_0}(p)$. In fact, since $h_\theta(q) \rightarrow +\infty$ as $q \rightarrow \theta_0$ we have

$$\bigcap_{q \in p_0\theta_0} J_{\theta_0}(q) = \{\theta_0\}.$$

Therefore,

$$\mu(J_{\theta_0}(q)) \rightarrow \mu(\{\theta_0\}) = 0.$$

It follows that there exists p on the ray from p_0 to θ_0 such that $\mu(J_{\theta_0}(p)) < 1$ and a compact set $K \subset X(\infty) \setminus J_{\theta_0}(p)$ such that $\mu(K) > 0$.

$$\begin{aligned} \int_{X(\infty)} h_\theta(q) d\mu(\theta) &= \int_{J_{\theta_0}(q)} h_\theta(q) d\mu(\theta) + \int_{X(\infty) \setminus J_{\theta_0}(q)} h_\theta(q) d\mu(\theta) \\ &\geq \int_{J_{\theta_0}(q)} h_\theta(q) d\mu(\theta) + \int_K h_\theta(q) d\mu(\theta), \end{aligned}$$

for q between p and θ_0 . Since for every $\theta \in K$, $h_\theta(p) \geq C > 0$, we have

$$\begin{aligned} h_\mu(q) = \int_{X(\infty)} h_\theta(q) d\mu(\theta) &\geq \frac{d(q, p_0)}{d(p, p_0)} C \mu(K) + \frac{d(q, p_0)}{d(p, p_0)} \int_{J_{\theta_0}(q)} h_\theta(q) d\mu(\theta) \\ &\geq \frac{d(q, p_0)}{d(p, p_0)} \left(C \mu(K) - \sup_{\theta \in X(\infty)} |h_\theta(p)| \mu(J_{\theta_0}(q)) \right). \end{aligned}$$

Since $\mu(J_{\theta_0}(q)) \rightarrow 0$ and $d(q, p_0) \rightarrow \infty$ as $q \rightarrow \theta_0$, the claim follows. \square

It should be noticed that the barycenter of a measure λ , being a critical point of h_λ , is defined implicitly by the equation

$$\int_{X(\infty)} dh_\theta(\cdot) d\lambda(\theta) = 0,$$

which is independent of the choice of p_0 , since changing p_0 changes h_λ by a constant.

4 The natural map

Let (M, g) and (N, g_0) be two n -dimensional compact, negatively curved manifolds. We assume that they are homotopically equivalent, i.e. there exist two continuous maps $f : M \rightarrow N$ and $h : N \rightarrow M$ such that $f \circ h$ is homotopic to the identity map of N and $h \circ f$ is homotopic to the identity of M . Since M and N are $K(\pi, 1)$ spaces, this hypothesis is equivalent to their fundamental groups being isomorphic as abstract groups. (See [2].)

We construct in this section a smooth map $F : M \rightarrow N$, the “natural map,” which will be the candidate for a homothety, when the conditions

of the main theorem are satisfied. The construction relies on a number of classical facts concerning manifolds of negative curvature, which can be found in complete detail in, say, [2].

Let M be a compact manifold of negative sectional curvature. Let X be the universal covering of M and $\Gamma \subset \text{Iso}(X)$ the group of deck transformations. Define $m_{p_0} = \sum_{\gamma \in \Gamma} \delta_{\gamma(p_0)}$, where δ_p is the Dirac measure concentrate at p . Then

$$\int_X e^{-sd(p,q)} dm_{p_0}(q) = \sum_{\gamma \in \Gamma} e^{-sd(p,\gamma(p_0))} =: g_s(p, p_0).$$

Lemma 4.1 *The critical exponent δ for the measure m_{p_0} is equal to the entropy $h(g)$.*

Proof. With the above notation, we recall that

$$\delta := \inf\{s \in [0, \infty] : \sum_{\gamma \in \Gamma} e^{-sd(p,\gamma(p_0))} < \infty\}.$$

Define

$$S_k := \# \left(\Gamma p_0 \cap (B_p(k + \frac{1}{2}) - B_p(k - \frac{1}{2})) \right).$$

Then

$$\begin{aligned} S_k &\sim c \text{Vol}(B_p(k + \frac{1}{2}) - B_p(k - \frac{1}{2})) \\ &= c V_p^g(k + \frac{1}{2}) \left(1 - \frac{V_p^g(k - \frac{1}{2})}{V_p^g(k + \frac{1}{2})} \right) \\ &\sim c(1 - e^{-h(g)}) V_p^g(k + \frac{1}{2}). \end{aligned}$$

On the other hand,

$$\begin{aligned} g_s(p, p_0) &\sim \sum_{k=0}^{\infty} S_k e^{-sk} \\ &= \sum_{k=0}^{\infty} e^{[\frac{\ln S_k}{k} - s]k}. \end{aligned}$$

Therefore, $\delta = \limsup_{k \rightarrow \infty} \frac{\ln S_k}{k} = h(g)$. □

We denote by ν_p the Patterson-Sullivan measures obtained as in section 2, and define

$$\mu_p = \frac{\nu_p}{\nu_p(X(\infty))},$$

a probability measure for each p . Γ -equivariance of $p \mapsto \nu_p$ implies that $p \mapsto \mu_p$ is also Γ -equivariant.

We now proceed to the construction of the natural map F .

1st. step. (Cf. [2, p. 84]) If M and N are homotopically equivalent, one can lift the maps f and h to a map between the universal covers X and Y of M and N , resp., in such a way that

$$\tilde{f}(\gamma(x)) = \rho(\gamma)\tilde{f}(x)$$

for all $x \in X$ and γ in the group of deck transformations $\pi_1(M)$. Here, ρ is the isomorphism between $\pi_1(M)$ and $\pi_1(N)$ induced by f . Moreover, by regularization, \tilde{f} and \tilde{h} can be taken to be C^1 maps. One can then show that \tilde{f} is a quasi-isometry between X and Y (cf. [2, p. 86]); here, compactness of M and N is essential. Finally, a quasi-isometry gives rise to an homeomorphism between the boundaries at infinity

$$\bar{f} : X(\infty) \rightarrow Y(\infty),$$

satisfying also the equivariance property

$$\bar{f} \circ \gamma = \rho(\gamma) \circ \bar{f},$$

where the action of the fundamental group $\pi_1(M)$ on X (resp., $\pi_1(N)$ on Y) is extended trivially to an action on $X(\infty)$ (resp. $Y(\infty)$.)

2nd step. The Patterson-Sullivan measure gives an equivariant map $p \mapsto \mu_p$ from X to the space of probability measures $\mathcal{M}_1(X(\infty))$ on the ideal boundary of X . For each p , μ_p has no atoms. We can now push forward each measure μ_p by the continuous map \bar{f} , thereby constructing a map

$$X \rightarrow \mathcal{M}_1(X(\infty)) \quad p \mapsto \bar{f}_*(\mu_p).$$

The equivariance property of \bar{f} under the action of $\pi_1(M)$ on $X(\infty)$, and on $Y(\infty)$ via the isomorphism ρ , shows that $p \mapsto \bar{f}_*(\mu_p)$ is equivariant with respect to the actions of $\pi_1(M)$ on X , and on $\mathcal{M}_1(Y(\infty))$ via ρ . Finally, since \bar{f} is a homeomorphism, the measures $\bar{f}_*(\mu_p)$ are well defined and have no atoms.

3rd step. We can now define the map \tilde{F} by

$$\tilde{F}(p) := \text{bar}(\bar{f}_*(\mu_p)).$$

It clearly satisfies the equivariance relation

$$\tilde{F}(\gamma(p)) = \rho(\gamma)\tilde{F}(p),$$

thus giving rise to a map $F : M \rightarrow N$. Its regularity will be studied in the next section. Since the map F induces the isomorphism ρ between the fundamental groups, it is homotopic to f .

Remark that \bar{f} is only required to be continuous. In fact, the only property of \bar{f} needed to prove the regularity of F is that

$$\bar{f}_* : \mathcal{M}_1(\mathbb{Y}(\infty)) \rightarrow \mathcal{M}_1(X(\infty))$$

exists, is linear and sends nonatomic measures to nonatomic measures.

5 Proof of the main theorem

In this section we prove the following special case of the main theorem.

Theorem 5.1 *Let (M, g) and (N, g_0) be two compact, negatively curved Riemannian n -manifold, such that (N, g_0) is locally symmetric. Suppose that M and N are homotopically equivalent. Then, if $n \geq 3$, we have*

1. $[h(g)]^n \text{Vol}(M, g) \geq [h(g_0)]^n \text{Vol}(N, g_0)$.
2. Equalities $h(g) = h(g_0)$ and $\text{Vol}(M, g) = \text{Vol}(N, g_0)$ occur if and only if (M, g) and (N, g_0) are isometric.

Most of the effort goes into showing the next proposition.

Proposition 5.2 *The natural map F is C^1 (at least). Furthermore, one has*

1. $|\text{Jac}F(p)| \leq \left(\frac{h(g)}{h(g_0)}\right)^n$ for all $p \in M$.
2. If for some $p \in M$, $|\text{Jac}F(p)| = \left(\frac{h(g)}{h(g_0)}\right)^n$, then the differential dF_p of F at p is a homothety (of ratio $\frac{h(g)}{h(g_0)}$).

To see how the above theorem follows from the proposition, we recall that F is a homotopy equivalence and so has degree one. Let ω_0 be the volume form of the (oriented) manifold (N, g_0) and ω the volume form of (M, g) . Then,

$$\int_M F^* \omega_0 = \int_N \omega_0 = \text{vol}(N, g_0)$$

and inequality (1) of the proposition gives

$$\begin{aligned} \text{Vol}(N, g_0) &\leq \int_M |F^* \omega_0| = \int_N |(\text{Jac} F) \omega| \\ &\leq \left(\frac{h(g)}{h(g_0)} \right)^n \int_M \omega \\ &\leq \left(\frac{h(g)}{h(g_0)} \right)^n \text{Vol}(M, g). \end{aligned}$$

This proves the first part of the Theorem. In the equality case then

$$|(\text{Jac} F)(p)| = \left(\frac{h(g)}{h(g_0)} \right)^n = 1,$$

for all $p \in M$ and hence dF_p is a homothety of ratio 1, i.e. and isometry for all $p \in M$.

We now proceed to the proof of the proposition. The same notation F will be used for a lift of $F : M \rightarrow N$ to the universal coverings X and Y . The inequalities of the proposition will be shown for such lift, but since they are pointwise, they will also hold for the map between the compact manifolds. Denote by B_θ^0 (resp. B_θ) the Busemann function of (Y, g_0) (resp. (X, g)). Let μ_p , $p \in X$ be the family of Patterson-Sullivan measures on $Y(\infty)$. Then the natural map can be defined implicitly by the equation

$$\int_{Y(\infty)} (dB_\theta^0)_{F(p)}(\cdot) d(\bar{f}_* \mu_p)(\theta) = 0,$$

a vector-valued equation. Equivalently, transferring the integral to the boundary of X , we have

$$\int_{X(\infty)} (dB_{\bar{f}(\alpha)}^0)_{F(p)}(\cdot) e^{-h(g)B_\alpha(p)} d(\mu_{p_0})(\alpha) = 0.$$

(We are using here the formula $\int_{Y(\infty)} l d(\bar{f}_* \mu) = \int_{X(\infty)} l \circ \bar{f} d\mu$.)

Choose a frame $\{e_i(q)\}_{i=1}^n$ of $T_q Y$ depending smoothly on q and define the function

$$G = (G_1, \dots, G_n) : Y \times X \rightarrow \mathbb{R}^n$$

such that

$$G_i(q, p) := \int_{X(\infty)} (dB_{\tilde{f}(\alpha)}^0)_q(e_i(q)) e^{-h(g)B_\alpha(p)} d\mu_0(\alpha).$$

Then the above integral, which implicitly defines F , takes the form

$$G(\bar{F}(p), p) = 0.$$

Since the Busemann functions B_θ^0 and B_α are smooth and $X(\infty)$ is compact, it is not difficult to see that G is smooth. Then the proof of the fact that F is C^1 is a simple application of the implicit function theorem. In fact, if we denote $\nabla_v e_i = \sum_j \eta_{ij}(v) e_j$, for $v \in TX$, the differential of G with respect to q becomes

$$(\partial_1 G_i)_{(q,p)} v = \sum_j \eta_{ij} G_j(q, p) + \int_{X(\infty)} (\nabla_v dB_{\tilde{f}(\alpha)}^0)(e_i(q)) e^{-h(g)B_\alpha(p)} d\mu_0(\alpha).$$

When $G(q, p) = 0$ the above reduces to

$$(\partial_1 G_i)_{(q,p)} v = \int_{X(\infty)} (\nabla_v dB_{\tilde{f}(\alpha)}^0)(e_i(q)) d\lambda_p(\alpha),$$

where λ_p is a nonatomic measure on $X(\infty)$. But we know that the latter integral is the Hessian of

$$B_{\lambda_p}^0 = \int_{X(\infty)} B_{\tilde{f}(\alpha)}^0 d\lambda_p,$$

which is positive definite, hence $\partial_1 G$ is invertible. Therefore the implicit function theorem implies that F is C^1 .

The implicit function theorem also gives a formula for the differential of F , as follows. Differentiating $G(F(p), p) = 0$ with respect to p , we get, for all $v \in T_p X$ and $u \in T_{F(p)} Y$:

$$\begin{aligned} 0 &= \int_{X(\infty)} (\nabla_{dF_p(v)} dB_{\tilde{f}(\alpha)}^0)_{F(p)}(u) d\mu_p(\alpha) \\ &\quad - h(g) \int_{X(\infty)} (dB_{\tilde{f}(\alpha)}^0)_{F(p)}(u) dB_\alpha(v) d\mu_p(\alpha). \end{aligned}$$

The above equality is to be understood as an equality between bilinear forms.

Define the following quadratic forms, K and H , on $T_{F(p)}Y$:

$$\begin{aligned} g_0(K_{F(p)}(u), u) &:= \int_{Y(\infty)} (\nabla_u dB_\theta^0)_{F(p)}(u) d(\bar{f}_* \mu_p)(\theta) \\ g_0(H_{F(p)}(u), u) &:= \int_{Y(\infty)} [(dB_\theta^0)_{F(p)}(u)]^2 d(\bar{f}_* \mu_p)(\theta). \end{aligned}$$

Combining the above formulas with the Cauchy-Schwarz inequality gives:

For all $u \in T_{F(p)}Y$ and $v \in T_pX$,

$$|g_0(K_{F(p)}(dF_p(v), u)| \leq h(g) (g_0(H_{F(p)}(u), u))^{\frac{1}{2}} \left(\int_{X(\infty)} [(dB_\alpha)_p(v)]^2 d\mu_p(\alpha) \right)^{\frac{1}{2}}.$$

It should be noticed that the symmetric endomorphism K is invertible since the bilinear form $g_0(K \cdot, \cdot)$ is the Hessian of the strictly convex function \mathcal{B} introduced before. This is what will allow us to compute the Jacobian of F . We recall that the Jacobian of F is the determinant of dF_p computed with respect to orthonormal basis of (T_pX, g) and $(T_{F(p)}Y, g_0)$.

Lemma 5.3 *With the above notations, we have:*

$$|\text{Jac}F(p)| \leq \frac{h^n(g)(\det H)^{1/2}}{n^{n/2} \det K}.$$

Proof. If the rank of dF_p is not maximal, the Jacobian determinant of F at p is zero and the inequality is trivially satisfied. We can therefore assume that dF_p is invertible. Let $\{u_i\}$ be an orthonormal basis of $T_{F(p)}Y$ which diagonalizes the endomorphism H . Define

$$v'_i = (K \circ dF_p)^{-1}(u_i).$$

(Recall that K is also invertible.) Applying the orthonormalization process to v'_i , we obtain a basis $\{v_i\}$ of T_pX . The matrix of $K \circ dF_p$ written in the basis $\{v_i\}$ and $\{u_i\}$ is then triangular, so that

$$\det(K \circ dF_p) = (\det K)(\text{Jac}F(y)) = \prod_{i=0}^n \langle K(dF_y(v_i)), u_i \rangle_0.$$

Here we identify endomorphisms with matrices using the basis involved. The previous inequality then gives

$$(\det K)|\text{Jac}F(y)| \leq h^n(g) \prod_{i=1}^n \langle H u_i, u_i \rangle_0^{1/2} \prod_{i=1}^n \left(\int_{X(\infty)} [(dB_\alpha)_p(v_i)]^2 d\mu_p(\alpha) \right)^{1/2}.$$

Using the fact that

$$\left(\prod_{i=1}^n a_i\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n a_i$$

and that

$$\sum_{i=1}^n [(dB_\alpha)_p(v_i)]^2 = \|(dB_\alpha)_p\|_g^2 = 1,$$

we obtain:

$$\begin{aligned} \prod_{i=1}^n \left(\int_{X(\infty)} [(dB_\alpha)_p(v_i)]^2 d\mu_p(\alpha) \right)^{1/2} &\leq \left(\sum_{i=1}^n \int_{X(\infty)} [(dB_\alpha)_p(v_i)]^2 d\mu_p(\alpha) \right)^{n/2} \\ &\leq \frac{1}{n^{n/2}}. \end{aligned}$$

The desired inequality now follows. \square

Recall now that if g_0 has constant sectional curvature -1 ,

$$\nabla dB^0(\cdot, \cdot) = g_0(\cdot, \cdot) - dB^0(\cdot)dB^0(\cdot),$$

which gives after integration

$$K = I - H.$$

It should also be noticed that

$$\text{trace}(H) = 1,$$

since

$$\begin{aligned} \text{trace}(H) &= \sum_{i=1}^n \langle H_{F(p)}(u_i), u_i \rangle_0 \\ &= \int_{X(\infty)} \sum_{i=1}^n [(dB_\theta^0)_{F(p)}(u_i)]^2 d(\bar{f}_* \mu_p)(\theta) \\ &= 1. \end{aligned}$$

The proposition now follows from the above results.

We recall that if $n \geq 3$,

$$\frac{\det(H)}{[\det(I - H)]^2} \leq \left[\frac{n}{(n-1)^2} \right]^n$$

and equality holds if and only if

$$H = \frac{1}{n}I.$$

We consider now the equality case, i.e. we suppose that

$$|\text{Jac}F(p)| = \left(\frac{h(g)}{h(g_0)} \right)^n.$$

Then for all $p \in X$

$$H_{F(p)} = \frac{1}{n}I \text{ and } K_{F(p)} = \frac{(n-1)}{n}I = \frac{h(g_0)}{n}I.$$

It follows that

$$|\langle dF_p(v), u \rangle_0| \leq n^{\frac{1}{2}} \frac{h(g)}{h(g_0)} \|u\|_0 \left(\int_{X(\infty)} [(dB_\alpha)_p(v)]^2 d\mu_p(\alpha) \right)^{\frac{1}{2}}$$

for all $v \in T_pX$ and $u \in T_{F(p)}Y$. By taking the supremum in $u \in T_{F(p)}Y$ such that $\|u\|_0 = 1$, one gets

$$\|dF_p(v)\|_0 \leq n^{\frac{1}{2}} \frac{h(g)}{h(g_0)} \left(\int_{X(\infty)} [(dB_\alpha)_p(v)]^2 d\mu_p(\alpha) \right)^{\frac{1}{2}}$$

for all $v \in T_pX$. Let L be the endomorphism of T_pX defined by

$$L = (dF_p)^* \circ (dF_p)$$

and $\{v_i\}$ a g -orthonormal basis of T_pX . Then we have

$$\begin{aligned} \text{trace}(L) &= \sum_{i=1}^n \langle Lv_i, v_i \rangle_g \\ &= \sum_{i=1}^n \langle dF_p(v_i), dF_p(v_i) \rangle_0 \\ &\leq n \left(\frac{h(g)}{h(g_0)} \right)^2, \end{aligned}$$

where we have again used the fact that $\|dB\|_g = 1$.

We have now

$$\left(\frac{h(g)}{h(g_0)} \right)^{2n} = |\text{Jac}F(p)|^2 = \det L \leq \left(\frac{1}{n} \text{trace} L \right)^n \leq \left(\frac{h(g)}{h(g_0)} \right)^{2n}.$$

Therefore the determinant of L is $(\frac{1}{n}\text{trace}L)^n$ and

$$L = \left(\frac{h(g)}{h(g_0)} \right)^2 I.$$

This precisely means that dF_p is an isometry composed with a homothety of ratio $h(g)/h(g_0)$.

References

- [1] W. Ballmann, M. Gromov, V. Schroeder. *Manifolds of nonpositive curvature*, Progress in Mathematics, Vol. 61, Birkhäuser, 1985.
- [2] R. Benedetti, C. Petronio. *Lectures on Hyperbolic Geometry*, Universitext, Springer-Verlag, 1991.
- [3] G. Besson, G. Courtois, S. Gallot. *A simple proof of the rigidity and minimal entropy theorems*, preprint, 1995.
- [4] G. Besson, G. Courtois, S. Gallot. *Entropies et rigidités des espaces localement symétriques de courbure strictement négative*, to appear in GAFA, Vol. 5, No. X (1995).
- [5] M. Burger, S. Mozes. *CAT(-1) spaces, divergence groups and their commensurators*, preprint, 1995.
- [6] M. Gromov. *Volume and bounded cohomology*, Publ. Math. Inst. Hautes Étud. Sci. **56** (1981), 213-307.
- [7] M. Gromov. *Filling Riemannian Manifolds*, J. Diff. Geom. **18** (1983) 1-147.
- [8] A. Katok. *Entropy and closed geodesics*, Ergodic Theory and Dynamical Systems, **2** (1982) 339-367.
- [9] A. Katok and B. Hasselblatt. *Introduction to the modern theory of Dynamical Systems*, Cambridge UP, 1995.
- [10] W. Klingenberg. *Riemannian Geometry*, de Gruyter studies in Math. **1**, 1982.
- [11] A. Manning. *Topological entropy for geodesic flows*, Ann. Math. **110** (1970), 567-573.

- [12] P. Nicholls. *The ergodic theory of discrete groups*, London Mathematical Society Lecture Notes Series no.143, Cambridge University Press, 1989.
- [13] S. Patterson. *The limit set of a fuchsian group*, Acta Math. **136** (1976), 241-273.

Renato Feres
Department of Mathematics
Washington University
St. Louis, Missouri 63130
feres@math.wustl.edu