Commun. Math. Phys. 208, 107 – 123 (1999)

Communications in
Mathematical
Physics

A Multiplicative Ergodic Theorem and Nonpositively Curved Spaces

Anders Karlsson*, Gregory A. Margulis**

Department of Mathematics, Yale University, New Haven, CT 06520, USA.
E-mail: andersk@math.yale.edu; margulis@math.yale.edu

Received: 27 April 1999 / Accepted: 25 May 1999

Abstract: We study integrable cocycles $u(n, x)$ over an ergodic measure preserving transformation that take values in a semigroup of nonexpanding maps of a nonpositively curved space Y, e.g. a Cartan–Hadamard space or a uniformly convex Banach space. It is proved that for any $y \in Y$ and almost all x, there exist $A \geq 0$ and a unique geodesic ray $\gamma(t, x)$ in Y starting at y such that

$$\lim_{n\to\infty} \frac{1}{n} d(\gamma(An, x), u(n, x)y) = 0.$$

In the case where Y is the symmetric space $GL_N(\mathbb{R})/O_N(\mathbb{R})$ and the cocycles take values in $GL_N(\mathbb{R})$, this is equivalent to the multiplicative ergodic theorem of Oseledec.

Two applications are also described. The first concerns the determination of Poisson boundaries and the second concerns Hilbert-Schmidt operators.

1. Introduction

Let (X, μ) be a measure space with $\mu(X) = 1$ and let $L : X \to X$ be a measure preserving transformation. Birkhoff's pointwise ergodic theorem asserts that the ergodic averages of a function $f \in L^1(\mu)$,

$$\frac{1}{n} \sum_{k=0}^{n-1} f(L^k x),$$

converge for μ-a.e. x to an L-invariant function $\bar{f} \in L^1(\mu)$ when $n \to \infty$.

Two important extensions of this theorem are the subadditive ergodic theorem of Kingman [Ki] and the multiplicative ergodic theorem of Oseledec [O]. Both theorems

* Supported in part by the Göran Gustafsson Foundation.
** Supported in part by NSF Grant DMS-9800607.

have numerous applications and since the original proofs were published several alternative proofs of these theorems have appeared. Let us first recall Kingman's theorem.

Let $a : \mathbb{N} \times X \to \mathbb{R} \cup \{-\infty\}$ be a subadditive (measurable) cocycle, that is

$$a(n+m, x) \leq a(n, L^m x) + a(m, x)$$

for $n, m \geq 1$ and $x \in X$. Assume that

$$\int_X a^+(1, x) d\mu(x) < \infty,$$

where $a^+(1, x) = \max\{0, a(1, x)\}$. Then the subadditive ergodic theorem asserts that there is an L-invariant measurable function $\overline{a} : X \to \mathbb{R} \cup \{-\infty\}$ such that

$$\lim_{n\to\infty} \frac{1}{n} a(n, x) = \overline{a}(x)$$

for μ-a.e. x.

This result generalizes Birkhoff's theorem, because

$$a(n, x) := \sum_{k=0}^{n-1} f(L^k x)$$

is a subadditive (in fact additive) cocycle.

The multiplicative ergodic theorem of Oseledec is an extension of Birkhoff's theorem to products of matrices. Let $A : X \to GL_N(\mathbb{R})$ be a measurable map and define the (multiplicative) cocycle

$$A(n, x) = A(L^{n-1}x) \cdots A(x).$$

Assume that

$$\int_X \log^+ ||A(x)|| d\mu(x) < \infty \text{ and } \int_X \log^+ ||A^{-1}(x)|| d\mu(x) < \infty,$$

where $\log^+ a = \max\{0, \log a\}$. Then the theorem of Oseledec asserts that for μ-a.e. x the sequence $A(n, x)$ is Lyapunov regular, which by definition means that there is a filtration of subspaces

$$\{0\} = V_0^x \subsetneq V_1^x \subsetneq \ldots \subsetneq V_{s(x)}^x = \mathbb{R}^N$$

and numbers $\lambda_1(x) < \ldots < \lambda_{s(x)}(x)$ such that for any $v \in V_i^x \setminus V_{i-1}^x$,

$$\lim_{n\to\infty} \frac{1}{n} \log ||A(n, x)v|| = \lambda_i(x)$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log |\det A(n, x)| = \sum_{i=1}^{s(x)} \lambda_i(x)(\dim V_i^x - \dim V_{i-1}^x).$$

Let W_i^x be the orthogonal complement of V_{i-1}^x in V_i^x and define a positive definite matrix $\Lambda(x)$ by requiring that $\Lambda(x)w = e^{\lambda_i(x)}w$ for any $w \in W_i^x$, $1 \leq i \leq s(x)$. The content of this theorem is that in a certain sense, $A(n, x)$ behave asymptotically like the iterates $\Lambda(x)^n$.

The Lyapunov regularity is also easily seen to be equivalent to the statement that there exists a positive definite symmetric matrix $\Lambda = \Lambda(x)$ such that

$$\frac{1}{n} \log ||A_n \Lambda^{-n}|| \to 0 \text{ and } \frac{1}{n} \log ||\Lambda^n A_n^{-1}|| \to 0, \tag{1.1}$$

where A_n denotes $A(n, x)$.

Consider the symmetric space $Y = GL_N(\mathbb{R})/O_N(\mathbb{R})$ and let $y = O_N(\mathbb{R})$. Let g be an element in $GL_N(\mathbb{R})$ and let μ_i denote the eigenvalues of $(gg^t)^{1/2}$. The distance in Y between y and gy is

$$d(y, gy) = \left(\sum_{i=1}^{N} (\log \mu_i)^2 \right)^{1/2}. \tag{1.2}$$

Recall also that geodesics starting at y are of the form $\gamma(t) = e^{tH} y$, where H is a symmetric matrix. Let $\Lambda = e^H$ be some positive definite symmetric matrix. From (1.2) it follows that

$$\frac{1}{n} d(\Lambda^{-n} y, A_n^{-1} y) \to 0, \tag{1.3}$$

is equivalent to (1.1).

Hence the Lyapunov regularity of $A(n, x)$ is equivalent to the geometric statement (1.3). For a discussion of this, see [Ka2]. In that paper, Kaimanovich obtained a complete geometric description of sequences $\{y_n\}$ of points in Y for which there are a geodesic ray γ and $A \geq 0$ such that the distance from y_n to $\gamma(An)$ grows sublinearly in n. This was done by taking advantage of the special structure of symmetric spaces of noncompact type, and using hyperbolic geometry. After that, applying the subadditive ergodic theorem, he could deduce (1.3).

The present paper studies the more general situation where the cocycles take values in a semigroup of semicontractions (e.g. isometries) of a uniformly convex, nonpositively curved in the sense of Busemann, complete metric space (Y, d). For definitions and examples, we refer to Sect. 3.

Note that the asymptotics of the iteration of one single semicontraction $\varphi : D \to D \subset Y$ is already nontrivial. For example, the case where D is a convex subset of a Hilbert space was studied by Pazy [P]. See also [Be] for this topic, which goes back to the work of Denjoy and Wolff on the iteration of an analytic map of the unit disk into itself.

In several proofs of Oseledec's theorem, the use of ergodic theory is reduced to the application of a standard theorem, that of Birkhoff or Kingman. In contrast, this reduction seems impossible to do for the proof of the multiplicative ergodic theorem given in this paper. Instead, we establish a different kind of "maximal ergodic inequality", Lemma 4.1. The arguments in the ergodic theoretic part of this paper are in the same spirit as those commonly used to establish the subadditive ergodic theorem. Note that, in the ergodic case, this theorem is here deduced as Corollary 4.3 of Proposition 4.2.

The paper is organized as follows. The section following this introduction, Sect. 2, provides a concise formulation of the main result. All the terminology used is explained in Sect. 3, which also contains one additional observation, Lemma 3.1. Section 4 proves the needed ergodic lemmas about subadditive cocycles (Proposition 4.2 and Corollary 4.3).

Section 5 gives the proof of the theorem. The final sections, Sects. 6 and 7, describe two applications.

2. Formulation of the Main Result

Let (Y, d) be a uniformly convex, complete metric space satisfying the Busemann non-positive curvature condition. Examples include CAT(0)-spaces and uniformly convex Banach spaces. Let S be a semigroup of semicontractions $D \to D$, where D is a nonempty subset of Y, and fix a point $y \in D$.

Furthermore, let (X, μ) be a measure space with $\mu(X) = 1$ and let $L : X \to X$ be an ergodic and measure preserving transformation. Given a measurable map $w : X \to S$, put

$$u(n, x) = w(x)w(Lx) \cdots w(L^{n-1}x) \tag{2.1}$$

and denote $u(n, x)y$ by $y_n(x)$. Assume that

$$\int_X d(y, w(x)y)d\mu(x) < \infty, \tag{2.2}$$

then the following "multiplicative ergodic theorem" holds.

Theorem 2.1. *For almost every x, the following limit exists:*

$$\lim_{n\to\infty} \frac{1}{n} d(y, y_n(x)) = A \tag{2.3}$$

and if $A > 0$, then for almost every x, there exists a unique geodesic ray $\gamma(\cdot, x)$ in Y starting at y such that

$$\lim_{n\to\infty} \frac{1}{n} d(\gamma(An, x), y_n(x)) = 0. \tag{2.4}$$

Remark 2.2. The existence of the limit (2.3) is well known. It is a standard consequence of the subadditive ergodic theorem, here Corollary 4.3. In the case $A > 0$, note that (2.4) implies that y_n converges to $[\gamma]$ in $Y \cup Y(\infty)$, where $Y(\infty)$ denotes the ideal boundary at infinity consisting of asymptote classes of rays.

Remark 2.3. Assume that $S = \Gamma$ is a discrete cocompact group of isometries of a Cartan-Hadamard manifold Y. Let P be the (time 1) Markov operator associated to a Γ-invariant Markov process on Y, with finite first moment and absolutely continuous transition probabilities. Take a P-stationary initial distribution on Y, then it is not difficult to construct a measure preserving system (X, μ, L) and a map $w : X \to \Gamma$, such that $u(n, x)y$ and the corresponding sample path at time n stay within a finite distance from each other for all n. The theorem then yields the result that for almost every sample path there is a geodesic ray such that the distance from the sample path to this geodesic grows sublinearly in n. In this context, we refer to Ballmann's paper [Ba1] for comparison.

Remark 2.4. There is also an "invertible" version of Oseledec's theorem, see [O], in which one gets the approximation by the powers of the same matrix at both $+\infty$ and $-\infty$, (the cocycle in question for negative n is $A(n, x) = A(1, L^n)^{-1} \cdots A(1, L^{-1})^{-1}$). In view of this result, one might wonder whether the analog statement for $u(n, x)$ is true in general, that is, is it true that there always exists a bi-infinite geodesic approximating both the backward and the forward orbit $u(n, x)y$ in the sense of Theorem 2.1? In general, however, the answer to this question is no. For example, let Y be the manifold $\mathbb{R} \times \mathbb{R}$ with Riemannian metric $(e^{-y} + C)^2 dx^2 + dy^2$. By some general results of Bishop and O'Neill concerning so-called warped products, the space Y is a Cartan-Hadamard manifold. Consider for $u(n, x)$ the powers of the parabolic isometry ϕ defined by $(x, y) \mapsto (x+1, y)$. Note that in this case the constant A in the theorem will equal C. If $C > 0$, then the forward and the backward orbit will converge to two different points on the ideal boundary of Y. These two limit points must be fixed by ϕ. Now assume that they can be connected by a geodesic in Y. Then, since the two endpoints are fixed by ϕ, the displacement of ϕ is semidecreasing in both directions along this geodesic, hence it is constant. This is impossible as ϕ is parabolic and Y has no parallel bi-infinite geodesics.

3. Geometric Preliminaries

General references for this section are [Ba2] and [J].

3.1. Let (Y, d) be a metric space. A continuous map $\gamma : I \to Y$, where I is an interval, is called a (unit speed minimizing) *geodesic*, if for any $s, t \in I$,

$$d(\gamma(s), \gamma(t)) = |s - t|.$$

A geodesic $\gamma : [0, \omega) \to Y$, such that $\lim_{t \to \omega} \gamma(t)$ does not exist, is called a *ray*. If (Y, d) is complete, then for any ray, $\omega = \infty$.

A point z is called a *midpoint* of x and y if

$$d(z, x) = d(z, y) = \frac{1}{2} d(x, y).$$

A metric space (Y, d) is called *convex* if any two points in Y have a midpoint. If a convex metric space (Y, d) is complete, then any two points can be joined by a geodesic.

A metric space (Y, d) is called *uniformly convex* if (Y, d) is convex and there is a strictly decreasing continuous function g on $[0, 1]$ with $g(0) = 1$, such that for any $x, y, w \in Y$ and midpoint z of x and y,

$$\frac{d(z, w)}{R} \le g\left(\frac{d(x, y)}{2R}\right),$$

where $R := \max\{d(x, w), d(y, w)\}$. See Fig. 1. An immediate consequence of this property is that midpoints are unique, and hence so are geodesics between any two points.

Spaces satisfying certain parallelogram inequalities, for example the L^p-spaces, $1 < p < \infty$, are uniformly convex, the original reference is [C]. For L^p, $p \ge 2$,

$$g(\varepsilon) = \left(1 - \varepsilon^p\right)^{1/p}$$

works in the definition. Further examples are Cartan-Hadamard manifolds (e.g. Euclidean spaces, hyperbolic spaces, and symmetric spaces of noncompact type such as

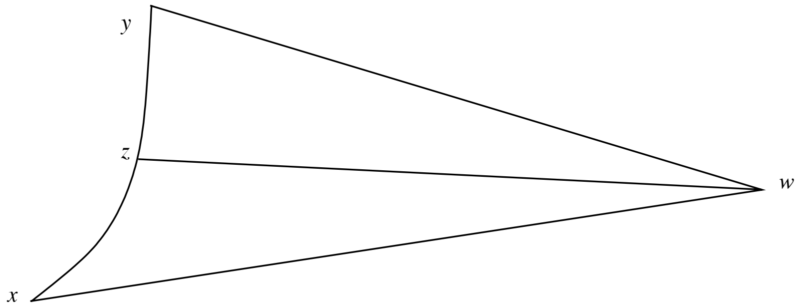

Fig. 1. The distance $d(z, w)$ is less than the maximum of $d(x, w)$ and $d(y, w)$

$GL_N(\mathbb{R})/O_N(\mathbb{R})$), or more generally CAT(0)- spaces (e.g. Euclidean buildings and $\mathbb{R}$-trees). For a general CAT(0)-space, g is as above with $p = 2$ and for trees one can also take $p = 1$. A Banach space is CAT(0) if and only if it is a Hilbert space.

A convex metric space (Y, d) is said to be *nonpositively curved in the sense of Busemann* if for any $x, y, z \in Y$ and any midpoints m_{xz} of x and z, and m_{yz} of y and z,

$$d(m_{xz}, m_{yz}) \leq \frac{1}{2} d(x, y). \tag{3.1}$$

Any uniformly convex Banach space, or more generally any strictly convex Banach space, as well as any CAT(0)-space satisfies Busemann's nonpositive curvature condition.

3.2. From now on, let (Y, d) be a uniformly convex, Busemann nonpositively curved, complete metric space.

It follows from the Busemann condition (3.1) that $t \to d(\gamma_1(t), \gamma_2(t))$ is a convex function for any two geodesics γ_1 and γ_2. In particular, for two rays γ_1 and γ_2 starting at y the function

$$t \to \frac{1}{t} d(\gamma_1(t), \gamma_2(t)) \tag{3.2}$$

is semiincreasing.

Let γ_i be any sequence of rays starting at y and assume that $\{\gamma_i(R)\}_{i=1}^{\infty}$ is a Cauchy sequence for every R. By the completeness of (Y, d), we can for each R define $\gamma(R) = \lim \gamma_i(R)$. It is then immediate that γ is a ray starting at y and we say that γ_i converges to γ.

Lemma 3.1. *Let $x, y, z \in Y$ and assume that*

$$d(y, x) + d(x, z) \leq d(y, z) + \delta d(y, x), \tag{3.3}$$

where $\delta \in [0, 1]$. Let w be the point on the geodesic between y and z such that $d(y, w) = d(y, x)$, then

$$d(w, x) \leq f(\delta) d(y, x),$$

where f is a function such that $f(s) \to 0$ as $s \to 0$. See Fig. 2.

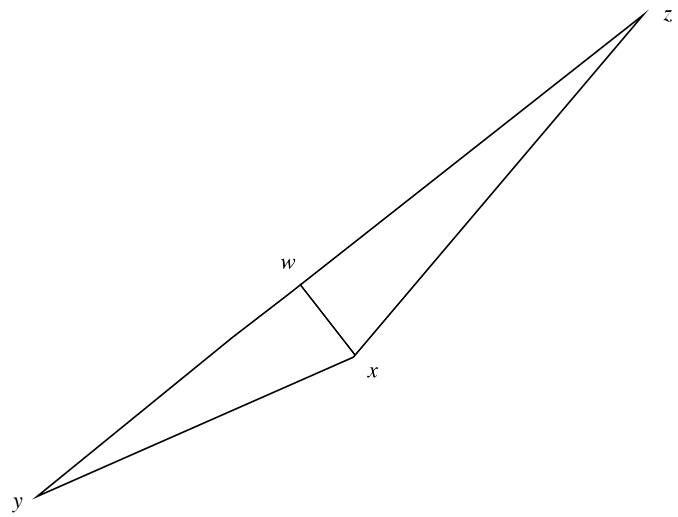

Fig. 2.

Proof. Let m be the midpoint of w and x. Uniform convexity implies that

$$d(m,z) \leq \max\{d(w,z), d(x,z)\}.$$

Since $d(w,z) = d(y,z) - d(y,x)$ by the definition of w and $d(x,z) \leq d(y,z) - d(y,x) + \delta d(y,x)$ by the inequality (3.3), we have that

$$d(m,z) \leq d(y,z) - d(y,x) + \delta d(y,x).$$

Hence it follows, by the triangle inequality, that

$$d(y,m) \geq d(y,x) - \delta d(y,x) = (1-\delta)R, \tag{3.4}$$

where $R := d(y,x) = \max\{d(y,x), d(y,w)\}$. Uniform convexity now gives us that

$$\frac{d(m,y)}{R} \leq g\left(\frac{d(w,x)}{2R}\right).$$

From the inequality (3.4) and since g is decreasing we get

$$g^{-1}(1-\delta) \geq \frac{d(w,x)}{2R}.$$

Recalling that $R = d(y,x)$ and letting $f(\delta) = 2g^{-1}(1-\delta)$, we have now obtained the desired conclusion. □

3.3. A *semicontraction* or *nonexpanding map* is a map $\varphi : D \to D$, where D is a subset of Y, such that

$$d(\varphi(y), \varphi(z)) \leq d(y,z)$$

for all $y, z \in D$. Any semigroup S of semicontractions is equipped with the Borel σ-algebra associated to the compact-open topology on S.

4. Ergodic Theoretic Part

Let (X, μ) be a measure space with $\mu(X) = 1$ and let $L : X \to X$ be a measure preserving transformation. Furthermore, let $a : \mathbb{N}\times X \to \mathbb{R}$ be a subadditive (measurable) cocycle, that is

$$a(n+m, x) \leq a(n, L^m x) + a(m, x) \tag{4.1}$$

for $n, m \in \mathbb{N}, x \in X$, (adopting the convention that $a(0, x) \equiv 0$). We will assume that the following integrability condition is satisfied:

$$\int_X a^+(1, x) d\mu(x) < \infty, \tag{4.2}$$

where $a^+(1, x) = \max\{0, a(1, x)\}$. For each n, let

$$a_n = \int_X a(n, x) d\mu(x). \tag{4.3}$$

It follows from (4.1) and (4.2) that $a_n \leq a_1 < \infty$, but it is possible that $a_n = -\infty$. Since L preserves μ, the subadditivity condition (4.1) implies that $a_{n+m} \leq a_n + a_m$ for every $n, m \in \mathbb{N}$. It is now an elementary fact, see for example [Kr, p. 36], that the limit

$$A := \lim_{n\to\infty} \frac{1}{n} a_n$$

exists and $A < \infty$.

Recall also the following observation of F. Riesz, which is proved by a simple induction, see [Bi, p. 27]. Let $c_1, c_2, ..., c_n$ be a finite sequence of real numbers. Call c_u a leader if at least one of the sums

$$c_u, c_u + c_{u+1}, ..., c_u + ... + c_n$$

is negative. Then the sum of the leaders is ≤ 0. (An empty sum is 0.)

Lemma 4.1. *Suppose that $A > 0$. Let E_1 be the set of x in X with the property that there are infinitely many n such that*

$$a(n, x) - a(n-k, L^k x) \geq 0$$

for all k, $1 \leq k \leq n$. Then $\mu(E_1) > 0$.

Proof. For every $i \in \mathbb{N}^+$ let us define a set

$$\Psi_i = \{x \in X | \exists k : 1 \leq k \leq i \text{ and } a(i, x) - a(i-k, L^k x) < 0\}$$

and a function

$$b_i(x) = a(i, x) - a(i-1, Lx).$$

It is clear that

$$\begin{aligned} &a(n, x) - a(n-k, L^k x) \\ &= b_n(x) + b_{n-1}(Lx) + ... + b_{n-k+1}(L^{k-1}x) \end{aligned} \tag{4.4}$$

and in particular

$$a(n,x) = b_n(x) + b_{n-1}(Lx) + \dots + b_1(L^{n-1}x). \tag{4.5}$$

In view of (4.4), if $L^k x \in \Psi_{n-k}$ then

$$b_{n-k}(L^k x) + \dots + b_{n-j}(L^j x) < 0$$

for some j, $k \le j \le n-1$. From this and F. Riesz's lemma about leaders (with $c_u := b_{n-u}(L^u x)$) we deduce that for every $x \in X$ and $n \in \mathbb{N}^+$,

$$\sum_{0\le k\le n-1,\, L^k x\in\Psi_{n-k}} b_{n-k}(L^k x) \le 0. \tag{4.6}$$

Using the L-invariance of μ, we get from the inequality (4.6) that

$$\begin{aligned}\sum_{j=1}^{n}\int_{\Psi_j} b_j(x)d\mu(x) = & \sum_{0\le k\le n-1}\int_{\Psi_{n-k}} b_{n-k}(x)d\mu(x)\\ = & \sum_{0\le k\le n-1}\int_{L^{-k}\Psi_{n-k}} b_{n-k}(L^k x)d\mu(x)\\ = & \int_X \sum_{0\le k\le n-1,\, L^k x\in\Psi_{n-k}} b_{n-k}(L^k x)d\mu(x) \le 0.\end{aligned} \tag{4.7}$$

On the other hand, in view of (4.3), (4.5), and the L-invariance of μ,

$$a_n = \int_X a(n,x)d\mu(x) = \sum_{j=1}^{n}\int_X b_j(x)d\mu(x). \tag{4.8}$$

Since $\lim a_n/n = A > 0$, there exists a number N such that

$$a_n > \frac{2A}{3}n \tag{4.9}$$

for all $n > N$.

Let Ψ_n^c denote the complement of Ψ_n in X. Then in view of (4.7), (4.8), (4.9), and the inequality $b_i(x) = a(i,x) - a(i-1,Lx) \le a(1,x) \le a^+(1,x)$, we have that

$$\sum_{j=1}^{n}\int_{\Psi_j^c} a^+(1,x)d\mu(x) \ge \sum_{j=1}^{n}\int_{\Psi_j^c} b_j(x)d\mu(x) > \frac{2A}{3}n \tag{4.10}$$

for all $n > N$. Let $f_n = \sum_{j=1}^{n}\chi_{\Psi_j^c}$, where χ_C denotes the characteristic function of a set $C \subset X$. Let $a_1^+ = \int_X a^+(1,x)d\mu(x)$ and

$$B_n = \{x \in X : n \ge f_n(x) > \frac{A}{3a_1^+}n\}.$$

Since

$$B_n^c = \{x \in X : \frac{A}{3a_1^+}n \ge f_n(x) \ge 0\},$$

we have that

$$\begin{aligned}\sum_{j=1}^{n}\int_{\Psi_j^c} a^+(1,x)d\mu(x) &= \int_X f_n(x)a^+(1,x)d\mu(x)\\ &= \int_{B_n} f_n(x)a^+(1,x)d\mu(x) + \int_{B_n^c} f_n(x)a^+(1,x)d\mu(x)\\ &\le n\int_{B_n} a^+(1,x)d\mu(x) + \frac{A}{3a_1^+}n\int_{B_n^c} a^+(1,x)d\mu(x)\\ &\le n\int_{B_n} a^+(1,x)d\mu(x) + \frac{A}{3}n.\end{aligned}$$

Combining this inequality and the inequality (4.10) we get that

$$\int_{B_n} a^+(1,x)d\mu(x) > \frac{A}{3} \tag{4.11}$$

for all $n > N$.

The condition (4.2) implies the existence of $\delta > 0$ such that

$$\int_C a^+(1,x)d\mu(x) < \frac{A}{3},$$

whenever $\mu(C) < \delta$. Hence it follows from (4.11) that $\mu(B_n) \ge \delta$ for every $n > N$. Let

$$C_n = \{x \in X : x \in \Psi_i^c \text{ for at least } \frac{A}{3a_1^+}n \text{ positive integers } i\},$$

so $B_n \subset C_n$ and $C_{n+1} \subset C_n$. Therefore, the measure of the set

$$\bigcap_{n\ge 1} C_n = \{x \in X : x \in \Psi_i^c \text{ for infinitely many } i\}$$

is greater than or equal to $\delta > 0$. Now recalling the definition of Ψ_i we get the desired statement. □

Proposition 4.2. *Suppose that L is ergodic and $A > -\infty$. For any $\varepsilon > 0$, let E_ε be the set of x in X for which there exist an integer $K = K(x)$ and infinitely many n such that*

$$a(n,x) - a(n-k, L^k x) \ge (A-\varepsilon)k$$

for all k, $K \le k \le n$. Let $E = \bigcap_{\varepsilon>0} E_\varepsilon$, then $\mu(E) = 1$.

Proof. For any $\varepsilon > 0$, let $c(n,x) = a(n,x) - (A-\varepsilon)n$. Then c is a subadditive cocycle and by the definition of A,

$$\lim_{n\to\infty}\frac{1}{n}\int_X c(n,x)d\mu = A - (A-\varepsilon) = \varepsilon > 0.$$

Note also that

$$a(n,x) - a(n-k, L^k x) \ge (A-\varepsilon)k$$

is equivalent to

$$c(n,x) - c(n-k, L^k x) \geq 0.$$

Hence Lemma 4.1 applied to c gives that $\mu(E_\varepsilon) > 0$.

By the subadditivity property (4.1),

$$a(n, L^l x) - a(n-k, L^{k+l}x) \geq a(n+l, x) - a((n+l)-(k+l), L^{k+l}x) - a(l, x).$$

It follows that $L^l E_\varepsilon \subset E_{2\varepsilon}$ for all $l \geq 0$ and by ergodicity we then get that $\mu(E_{2\varepsilon}) = 1$. Since this holds for every $\varepsilon > 0$ and $E_\varepsilon \subset E_{\varepsilon'}$, whenever $\varepsilon < \varepsilon'$, we have that $\mu(E) = 1$. □

Corollary 4.3 (Kingman). *Suppose that L is ergodic and $A > -\infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} a(n,x) = A$$

for almost every x.

Proof. Note that, by subadditivity, Proposition 4.2 implies that the set of x such that

$$\liminf_{k\to\infty} \frac{1}{k} a(k,x) \geq A - \varepsilon$$

for any $\varepsilon > 0$ has full measure. If $a(n,x)$ is an additive cocycle, then the a.e. convergence is immediate, since in this case the above proposition can also be applied to $-a(k,x)$.

In the case of a general subadditive cocycle $a(n,x)$, we can therefore subtract the additive cocycle

$$\sum_{i=0}^{n-1} a(1, L^i x)$$

from $a(n,x)$. This reduces the general case to the case of a nonpositive subadditive cocycle, that is $a(n,x) \leq 0$.

Fix an $\varepsilon > 0$ and take M such that

$$\frac{1}{M}\int_X a(M,x)d\mu(x) \leq A + \varepsilon \tag{4.12}$$

and let

$$a^M(n,x) = a(nM, x) - \sum_{i=0}^{n-1} a(M, L^{iM}x).$$

This $a^M(n,x)$ is again a nonpositive subadditive cocycle. From the proposition and the inequality (4.12), we have that

$$0 \geq \liminf_{n\to\infty} \frac{1}{nM} a^M(n,x) \geq -\varepsilon.$$

From this inequality, the nonpositivity and subadditivity of $a(n, x)$, the L-invariance and the convergence for additive cocycles, it follows that

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n}a(n,x) - \liminf_{n\to\infty} \frac{1}{n}a(n,x) &= \limsup_{n\to\infty} \frac{1}{nM}a(nM,x) - \liminf_{n\to\infty} \frac{1}{nM}a(nM,x) \\ &= \limsup_{n\to\infty} \frac{1}{nM}a^M(n,x) - \liminf_{n\to\infty} \frac{1}{nM}a^M(n,x) \\ &\leq -\liminf_{n\to\infty} \frac{1}{nM}a^M(n,x) \leq \varepsilon.\end{aligned}$$

Since this holds for any $\varepsilon > 0$, the corollary is established. For more details, consult [Kr, p. 37]. □

5. Proof of the Theorem

5.1. Here we adopt the notations in Sect. 2 and we let $a(n, x) = d(y, y_n(x))$. By the triangle inequality, the equality (2.1), and the semicontraction property,

$$\begin{aligned}d(y, y_{n+m}(x)) &\leq d(y, y_m(x)) + d(y_m(x), y_{n+m}(x)) \\ &= a(m,x) + d(u(m,x)y, u(m,x)u(n, L^m x)y) \\ &\leq a(m,x) + a(n, L^m x),\end{aligned}$$

hence a is a subadditive cocycle. Furthermore, by the assumption (2.2),

$$\int_X a^+(1,x)d\mu(x) = \int_X d(y, w(x)y)d\mu(x) < \infty,$$

which means that the basic integrability condition (4.2) of the cocycle a is satisfied. Corollary 4.3 (the subadditive ergodic theorem) then implies that

$$\lim_{n\to\infty} \frac{1}{n}d(y, y_n(x)) = A \geq 0 \tag{5.1}$$

for almost every $x \in X$.

5.2. Assume now that $A > 0$. Let E be the set defined as in Proposition 4.2 and consider an $x \in E$ such that (5.1) holds. From now on, x will frequently be suppressed in the notation.

For each $i > 0$, pick ε_i so small that

$$f(\delta_i) \leq 2^{-i},$$

where $\delta_i := 2\varepsilon_i/(A - \varepsilon_i)$ and f is the function appearing in the geometric lemma (3.1). This is possible, since $f(t) \to 0$ as $t \to 0$.

Proposition 4.2 and Corollary 4.3 give us that there are for any i an integer K_i and infinitely many n such that

$$a(n,x) - a(n-k, L^k x) \geq (A - \varepsilon_i)k \tag{5.2}$$

and

$$(A - \varepsilon_i)k \leq a(k, x) \leq (A + \varepsilon_i)k \tag{5.3}$$

for all $k,\ K_i \leq k \leq n$.

For each i, pick an integer n_i greater than both n_{i-1} and K_{i+1}, such that (5.2) and (5.3) hold. By adding the inequality (5.2) to the right inequality in (5.3), we get that for all $k,\ K_i \leq k \leq n_i$,

$$a(n_i, x) - a(n_i - k, L^k x) + (A + \varepsilon_i)k \geq (A - \varepsilon_i)k + a(k, x),$$

which simplified becomes

$$a(k, x) + a(n_i - k, L^k x) \leq a(n_i, x) + 2\varepsilon_i k.$$

From this, recalling the definition of a, the semicontractivity of $u(k, x)$, and the left inequality in (5.3), we get (note that at this point the order in which the maps $w(L^k x)$ are composed to form $u(n, x)$ is crucial)

$$\begin{aligned} d(y, y_k) + d(y_k, y_{n_i}) &\leq d(y, y_{n_i}) + 2\varepsilon_i k \\ &\leq d(y, y_{n_i}) + \frac{2\varepsilon_i}{A - \varepsilon_i} d(y, y_k). \end{aligned} \tag{5.4}$$

For each i, let γ_i be a ray from y passing through y_{n_i} and let $r_k = d(y, y_k)$. Applying the geometric lemma (3.1) to (5.4), we get that

$$d(\gamma_i(r_k), y_k) \leq f(\delta_i) r_k, \tag{5.5}$$

for all $k,\ K_i \leq k \leq n_i$.

5.3. We now show that $\{\gamma_i(R)\}$ is a Cauchy sequence for every $R > 0$. Fix $R > 0$. Since $K_{i+1} < n_i < n_{i+1}$, the inequality (5.5) implies that

$$\begin{aligned} d(\gamma_{i+1}(r_{n_i}), \gamma_i(r_{n_i})) &= d(\gamma_{i+1}(r_{n_i}), y_{n_i}) \\ &\leq f(\delta_{i+1}) r_{n_i}. \end{aligned}$$

For i large enough so that $r_{n_i} > R$, the convexity property (3.2) implies that

$$d(\gamma_{i+1}(R), \gamma_i(R)) \leq f(\delta_{i+1}) R,$$

which means, using the triangle inequality, that

$$d(\gamma_{i+m}(R), \gamma_i(R)) \leq \sum_{j=1}^{m} f(\delta_{i+j}) R \leq 2^{-i} R$$

for all $m > 0$. Hence $\{\gamma_i(R)\}$ is a Cauchy sequence and by the completeness of Y, γ_i converges to some ray γ, as $i \to \infty$.

5.4. It remains to show that

$$\lim_{k\to\infty}\frac{1}{k}d(\gamma(Ak), y_k) = 0.$$

For any k there is an i such that $K_i \leq k \leq n_i$ and by the triangle inequality

$$\begin{aligned} d(\gamma(Ak), y_k) &\leq d(\gamma(Ak), \gamma_i(Ak)) + d(\gamma_i(Ak), \gamma_i(r_k)) + d(\gamma_i(r_k), y_k) \\ &\leq 2^{-i}Ak + |Ak - r_k| + f(\delta_i)r_k \\ &\leq 2^{-i}Ak + \varepsilon_i k + f(\delta_i)(A+\varepsilon_i)k \leq (2^{-i+1}A + 2\varepsilon_i)k. \end{aligned}$$

It is then clear that

$$\limsup_{k\to\infty}\frac{1}{k}d(\gamma(Ak), y_k) \leq 0,$$

which shows (2.4). The uniqueness of γ is immediate from the convexity property (3.2) and so Theorem 2.1 is proved.

6. An Application to Random Walks and Boundary Theory

General references for this section are [F] and [Ka3].

Let Γ be a countable group acting by isometries on a uniformly convex, Busemann nonpositively curved, complete metric space (Y, d). Any isometry of Y also acts on the ideal boundary at infinity $Y(\infty)$, which consists of asymptote classes of geodesic rays.

Let ν be a probability measure on Γ and assume throughout this section that ν has finite first moment, that is

$$\sum_{g\in\Gamma} d(y, gy)\nu(g) < \infty.$$

Let (X, μ) be the product of $\mathbb{Z}$ copies of (Γ, ν) and let L be the shift transformation. This is a standard construction of an ergodic measure preserving system with $\mu(X) = 1$.

Let $w : X \to \Gamma$ be the projection onto the 0^{th} copy of Γ, so

$$w(\{g(i)\}_{i=-\infty}^{\infty}) = g(0),$$

and put as usual $u(n, x) = w(x)w(Lx)\cdots w(L^{n-1}x)$. This is sometimes called the right random walk determined by ν. Note that, in probabilistic language, the increments $\{w \circ L^k\}_{k=1}^{\infty}$ are independent, identically distributed random variables.

Theorem 2.1 (in the case $A > 0$) now provides a measurable map

$$\xi : X \to Y(\infty),$$

where $\xi(x) = [\gamma(x, .)]$. Since $u(n, Lx) = w(x)^{-1}u(n, x)$, the map ξ clearly has the following equivariance property:

$$\xi(Lx) = w(x)^{-1}\xi(x).$$

It follows that $(Y(\infty), \xi_*(\mu))$ is a ν-boundary for Γ. In the case $A = 0$, we set $\xi_*(X, \mu)$ to be the trivial ν-boundary for Γ.

Recall Kaimanovich's ray approximation criterion for the maximality of a boundary [Ka1, Theorem 3].

Theorem 6.1 (Kaimanovich). *Let (B, λ) be a ν-boundary of Γ and assume that ν has finite entropy $H(\nu) = -\sum \nu(g) \log \nu(g)$. Suppose that $\theta : \Gamma \to Z$ is a mapping into a metric space with metric d, and $\pi_n : B \to Z$ is a family of measurable mappings, and that there is a constant $C > 0$ such that*

$$\text{card}\{g \in \Gamma : d(z, \theta(g)) \leq N\} \leq e^{CN} \tag{6.1}$$

for all $z \in Z$ and $N \geq 1$. Let $b(x)$ denote the image in B of the sample path $\{u(n, x)\}$. If

$$\lim_{n\to\infty} \frac{1}{n} d(\pi_n(b(x)), \theta(u(n, x))) = 0$$

for almost every x, then the Poisson boundary of (Γ, ν) is isomorphic to (B, λ).

The following statement is now an immediate consequence.

Corollary 6.2. *Let Γ be a countable group acting on (Y, d) by isometries and let ν be a probability measure on Γ with finite first moment. Fix a point $y \in Y$ and assume that for some $C > 0$,*

$$\text{card}\{g \in \Gamma : d(y, gy) \leq N\} \leq e^{CN} \tag{6.2}$$

for all $N \geq 1$. Then the Poisson boundary of (Γ, ν) is isomorphic to $\xi_*(X, \mu)$.

Proof. Set $\theta(g) = gy$, $Z = Y$, $B = \xi(X)$, $\lambda = \xi_*(\mu)$, and $\pi_n(b(x)) = \gamma(An, x)$ using the notation of Theorem 2.1.

Since Γ acts by isometries it follows that

$$\text{card}\{g : d(z, gy) \leq N\} \leq \text{card}\{g : d(y, gy) \leq 2N\},$$

which ensures that condition (6.1) is satisfied. From this condition and the finiteness of the moment of ν, it follows that the entropy of ν is finite, see [Ka3]. □

Remark 6.3. When the group generated by suppν is nonamenable, the Poisson boundary is non-trivial, see [F], and so in particular $A > 0$. It is also known and not hard to show that the condition (6.2) is satisfied if Γ is a discrete subgroup of isometries of a locally compact Cartan-Hadamard manifold with sectional curvatures bounded from below.

Remark 6.4. Results on the determination of the Poisson boundary for various groups and measures have been obtained by many authors, see [Ka3]. Ballmann and Ledrappier in [BaLe] identified the Poisson boundary for cocompact lattices in rank 1 manifolds for nondegenerate measures with finite first moment and finite entropy (Kaimanovich was later able to replace the finite first moment with finite logarithmic moment, see [Ka3]). Note that their techniques are quite different from the methods in the present paper. Some of the ideas in [Ba1,BaLe], and [Ka3] go back to Furstenberg's work.

7. An Application to Hilbert-Schmidt Operators

Let $\mathcal{H}$ be a real Hilbert space and let $\mathcal{A}$ be the algebra of Hilbert-Schmidt operators $\mathcal{H} \to \mathcal{H}$, that is $a \in \mathcal{A}$ if

$$||a||_2^2 := \mathrm{tr}(aa^*) = \sum_i ||ae_i||^2 < \infty,$$

for some (hence any) orthonormal basis $\{e_i\}$ of $\mathcal{H}$. Recall that

$$< a, b >:= \mathrm{tr}(ab^*)$$

is an inner product on $\mathcal{A}$ and if $|| \cdot ||$ denotes the usual operator norm then

$$|| \cdot || \leq || \cdot ||_2. \tag{7.1}$$

It is a standard fact that $(\mathcal{A}, <, >)$ is a Hilbert space. Note also that the Cauchy-Schwarz inequality

$$(\mathrm{tr}(ab^*))^2 \leq \mathrm{tr}(aa^*)\mathrm{tr}(bb^*),$$

with $a = vw$, $b = wv$, where $v = v^*$ and $w = w^*$ yields

$$\mathrm{tr}(vwvw) \leq \mathrm{tr}(v^2w^2).$$

Now let $\mathrm{Sym} = \{a \in \mathcal{A} : a = a^*\}$ and $\mathrm{Pos} = \exp\{\mathrm{Sym}\} \subset I + \mathrm{Sym}$, where exp is the usual exponential map and I is the identity operator. Pos is an infinite dimensional Riemannian manifold with the metric

$$< v, w >_p:= \mathrm{tr}(p^{-1}vp^{-1}w),$$

$p \in \mathrm{Pos}$, $v, w \in \mathrm{Sym} \simeq T_p\mathrm{Pos}$. Let d be the associated distance function.

The arguments in [La, Ch. XII] show that (Pos, d) is a complete metric space satisfying the semi-parallelogram law and also that the operators $\exp\{\mathcal{A}\}$ act on Pos by isometries,

$$p \mapsto [\exp(a)]p := \exp(a)p\exp(a)^*.$$

Hence this is a situation in which Theorem 2.1 applies.

Corollary 7.1. *Let $u(n, x)$ be an integrable cocycle taking values in* $\exp\{\mathcal{A}\}$. *Then for almost every x there is an operator $\Lambda(x) = \exp(v(x))$, $v(x) \in$ Sym, such that*

$$\frac{1}{n}\left\|\log([\Lambda^{-n}(x)u(n, x)]I)\right\|_2 = \frac{1}{n}\left(\sum_i (\log \mu_i(n))^2\right)^{1/2} \to 0,$$

where $\mu_i(n)$ are the eigenvalues of $[\Lambda^{-n}(x)u(n, x)]I \in \mathrm{Pos}$.

The following Lyapunov regularity statement is a consequence of this corollary.

Let $\{f_i(x)\}$ be the orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of $\Lambda(x)$, so $\Lambda(x) f_i(x) = \exp(\lambda_i(x)) f_i(x)$. For $z = \sum_i z_i(x) f_i(x) \in \mathcal{H}$, let

$$\lambda_z(x) = \sup\{\lambda_i(x) : z_i(x) \neq 0\}.$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \log ||u(n,x)^{-1} z|| = -\lambda_z(x). \tag{7.2}$$

In [R], Ruelle obtained this type of multiplicative ergodic theorems for more general classes of operators. Note, however, that in the case of the Hilbert-Schmidt operators that we consider here, it is not clear that Corollary 7.1, which in infinite dimensions is a stronger statement than (7.2), can be proved by the methods in [R].

Acknowledgements. The authors would like to thank Vadim Kaimanovich for useful comments on an earlier version of this paper.

References

[Ba1] Ballmann, W.: On the Dirichlet problem at infinity for manifolds of nonpositive curvature. Forum Math. **1**, 201–213 (1989)

[Ba2] Ballmann, W.: *Lectures on Spaces of Nonpositive Curvature*. DMV-Seminar, Bd. 25. Basel, Boston, Berlin: Birkhäuser, 1995

[BaLe] Ballmann, W., Ledrappier, F.: The Poisson boundary for rank one manifolds and their cocompact lattices. Forum Math. **6**, 301–313 (1994)

[Be] Beardon, A.F.: Iteration of contractions and analytic maps. J. London Math. Soc. **41**, 141–150 (1990)

[Bi] Billingsley, P.:, *Ergodic Theory and Information*. New York: Wiley, 1965

[C] Clarkson, J.A.:, Uniformly convex spaces. Trans. Am. Math. Soc. **40**, 396–414 (1936)

[F] Furstenberg, H.: *Boundary theory and stochastic processes on homogeneous spaces*. Proc. Symp. Pure Math., Vol. **26**, Providence, RI: American Mathematical Society, 1973, pp. 193–229

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

[Ka1] Kaimanovich, V.A.: An entropy criterion for maximality of the boundary of random walks on discrete groups. Soviet Math. Dokl. **31**, 193–197 (1985)

[Ka2] Kaimanovich, V.A.: Lyapunov exponents, symmetric spaces and multiplicative ergodic theorem for semisimple Lie groups. J. Soviet Math. **47**, 2387–2398 (1989)

[Ka3] Kaimanovich, V.A.: The Poisson formula for groups with hyperbolic properties. Prépublication 97–13, Institut de Recherche Mathématique de Rennes, 1997

[Ki] Kingman, J.F.C.: The ergodic theory of subadditive stochastic processes. J. Roy. Statist. Soc. B **30**, 499–510 (1968)

[Kr] Krengel, U.: *Ergodic Theorems*. de Gruyter Stud. in Math., Vol. **6**. New York: de Gruyter, 1985

[La] Lang, S.: *Fundamentals of Differential Geometry*. New York: Springer-Verlag, 1999

[O] Oseledec, V.I.: A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems. Trans. Moscow Math. Soc. **19**, 197–231 (1968)

[P] Pazy, A.: Asymptotic behavior of contractions in Hilbert space. Israel J. Math. **9**, 235–240 (1971)

[R] Ruelle, D.: Characteristic exponents and invariant manifolds in Hilbert space. Ann. Math. **155**, 243–290 (1982)

Communicated by Ya. G. Sinai